KILLER ALERTS? PUBLIC HEALTH WARNINGS AND HEAT STROKE IN JAPAN*

Lester Lusher,[a] Tim Ruberg[b†]

[a]University of Pittsburgh and IZA

[b]University of Hohenheim

October 18, 2023

PRELIMINARY DRAFT, PLEASE DO NOT CITE/CIRCULATE

Abstract

In 2020, Japan introduced a comprehensive heat-health warning system where daily alerts were issued by region when forecasted wet bulb globe temperature (WBGT) exceeded a threshold ($33°C$). Utilizing plausibly exogenous region-day variation in the difference between actual and forecasted WBGT (i.e. forecasting errors), we find that the alerts led to a large and precisely estimated increase in heat stroke counts. Paired with data from Google Trends, Google Mobility Reports, and the population of ambulance records, we identify potential mechanisms, including increased reporting of heat stroke cases and "adverse" behavioral responses (e.g. people spending more time outdoors) when alerts were issued, while ruling out potential substitution in health diagnoses away from other sudden illnesses.

Keywords: heat stroke, climate change, warning effectiveness, avoidance behavior

JEL codes: D90, I12, I18, Q54

*We thank Mika Akesaka, Scott Carrell, Timothy Halliday, Shotaro Nakamura, Matt Naven, Yuki Nohara, Akira Sasahara, Hitoshi Shigeoka, Jeffrey Shrader, Manasvini Singh, Mirka Zvedelikova, and seminar participants at the University of Hohenheim, University of Tokyo, Osaka University, and Keio University for helpful comments. We thank the Fire and Disaster Management Agency (FDMA) for providing micro-level data. The views expressed in this paper are those of the authors and do not necessarily reflect the positions or policies of the organizations mentioned above. Any remaining errors are our own. Declarations of interest: none.

†Lester Lusher, Department of Economics, Pittsburgh, 230 South Bouquet Street, PA 15260, United States; IZA; Email: lester-lusher@pitt.edu. Tim Ruberg, University of Hohenheim, Department of Economics, Schloss Hohenheim, 70593 Stuttgart, Germany; Email: tim.ruberg@uni-hohenheim.de.

1 Introduction

In 2023, the world experienced one of the hottest summers ever recorded in human history. July 3, 2023, was the hottest day ever recorded, a record broken the very next day (and again on July 5 and July 6).[1] Extreme heat is the leading cause of weather-related death in the United States (more so than hurricanes, floods, and tornadoes, for example),[2] with the Los Angeles Times recently calling extreme heat the "quiet" health crisis.[3] Worldwide, extreme heat is a significant, and likely under-counted, cause of overall death (Ebi et al., 2021). Climate change has led to a drastic increase in average global warming, which has caused an increase in the incidence of record high temperatures and extreme heat events measured across weather stations (Robinson et al., 2021). Historically, single extreme heatwaves have claimed thousands of lives, like in the United States in 1995 (Palecki et al., 2001), Europe in 2003 and 2022 (D'Ippoliti et al., 2010; Ballester et al., 2023), and India in 2015 (Dodla et al., 2017), raising great concern for the implications of climate change in the present and immediate future.[4]

In this study, we investigate the impacts of a public warning system for extreme heat events on health and behavioral outcomes. Despite their potential importance and growing usage, relatively little is known about the efficacy of heat warning systems.[5] The context of our study is Japan. To our knowledge, Japan is one of only several high-income countries to have adopted a heat-health warning system.[6] Experts worldwide agree that there is a global need for heat-health warning systems, yet are severely lacking, even among high-income countries (Li et al., 2022). In the United States, a recent 2022 National Public Radio article dissected the insufficiencies of the current heatwave warning system.[7] In-progress examples of heat warning systems at the state level include the HeatRisk program covering a handful of western U.S. states, currently in a piloting phase,[8] and an impending program in California, which approved a bill in September 2022

[1]Source: PBS News Hour. Website: https://www.pbs.org/newshour/science/summer-of-record-breaking-heat-paints-story-of-a-warming-world-scientists-say, retrieved July 25, 2023.

[2]U.S. National Weather Service. Website: https://www.weather.gov/hazstat/, retrieved February 28, 2023.

[3]Los Angeles Times. Website: https://centerforhealthjournalism.org/2021/11/04/extreme-heat-overlooked-health-crisis-worth-covering-la-times-investigation-shows, retrieved March 1, 2023.

[4]See Martiello and Giacchi (2010) for a review of the relationship between high temperatures and mortality and morbidity outcomes.

[5]Small-scale and/or correlational studies from epidemiology assessing the potential health impacts from heat warnings include Weinberger et al. (2018), Benmarhnia et al. (2019), Vaidyanathan et al. (2019), and Weinberger et al. (2021).

[6]Other high-income countries to have adopted a heat-health warning system as of 2022 include France, the UK, and Spain (Kotharkar and Ghosh, 2022).

[7]For example, the primary variable for the U.S. warning system is the National Weather Service's "heat index," which underestimates the effect of extreme temperatures on the human body, thus underselling the true hazard from heat. Furthermore, only vague information is given regarding health hazards in response to heat from the U.S. warning system. National Public Radio, "Why heat wave warnings are falling short in the U.S." Website: https://www.npr.org/2022/09/13/1122491546/why-heat-wave-warnings-are-falling-short-in-the-u-s, retrieved March 1, 2023.

[8]National Weather Service, "NWS HeatRisk Prototype." Website: https://www.wrh.noaa.gov/wrh/heatrisk/, retrieved March 1, 2023.

requiring the state to develop a heat wave ranking system.[9] Japan piloted a heat-health warning system in 2020, then fully implemented the program nationally in 2021, granting us three years of observations with a warning system in place (2020–2022).

Additional advantages from the Japanese context include comprehensive hospitalization data and features of the warning system, which allow us to plausibly identify the causal impacts of the warning system on health outcomes. First, Japan provides extensive information for heat stroke outcomes by severity, demographic, and location, among other variables, whereas in the United States, hospitals and health care providers are not required to report heat-related illnesses to public health agencies.[10] Furthermore, in Japan, micro-level data exist for the population of ambulance dispatches, including various information on patient health outcomes and location.

For plausibly exogenous variation, Japan issues warnings across spatial areas ($n = 44$ regions) based on whether the *forecasted* wet bulb globe temperature (WBGT) exceeds a certain threshold (33.0$^{\circ}C$). Combined with data on whether a warning was issued for a region and *actual* WBGT, our econometric model effectively allows us to estimate the effect of a heat warning while flexibly controlling for local daily WBGT. In other words, our data include region-dates where an alert was issued but the actual WBGT was below 33.0$^{\circ}C$ (i.e. a false positive) *and* region-dates where an alert was not issued but the actual WBGT exceeded 33.0$^{\circ}C$ (i.e. a false negative). Thus, our econometric model separately disentangles the effects of a high heat warning from high heat itself using exogenous variation in forecasting errors.

We precisely estimate a substantial *increase* in heat stroke cases in response to a heat-health warning being issued. For our fully specified model, which flexibly controls for same-day and prior-day local WBGT (i.e. dummies for each WBGT level, rounded to the nearest tenth) as well as prefecture and date fixed effects, a heat-health warning increases the incidence of heat stroke hospitalizations by nearly 17%. The effects exist across the spectrum of severity, including "severe" cases which require multiple weeks of hospitalization. We further show the robustness of these results to various considerations, including additional controls for region-specific time trends and COVID-19 policies, alternative difference-in-differences estimators, estimating OLS vs. log-linear vs. negative binomial regression models, different sample restrictions, randomization inference, and intertemporal models. The result still appears even in a simple descriptive figure, where on average, for the exact same WBGT, region-days with a heat alert experience more heat strokes than region-days without a heat alert.

We then consider several potential mechanisms to explain these seemingly counter-intuitive results.

[9] California Legislative Information, "AB-2238 Extreme heat: statewide extreme heat ranking system." retrieved March 1, 2023.

[10] Centers for Disease Control and Prevention, "Picture of America Heat-Related Illness Fact Sheet," retrieved March 1, 2023

First, it could be that the health issues associated with extreme heat often go unidentified and/or unreported, and by raising the awareness of the hazards induced by high heat, the heat-health alert effectively encouraged people to report what were otherwise unidentified cases of heat stroke (i.e. an "increased reporting" effect). Second, it may be that cases of heat stroke are sometimes conflated with other sudden illnesses (such as actual stroke, heart attacks, and other cardiovascular issues), and when an alert is issued, there's effectively a "substitution" in health diagnoses away from other sudden illnesses and into heat stroke. Note that in both of these instances, the "actual" incidence of heat stroke is unknown (and possibly decreasing), it's just that "increased reporting" and/or "substituting" lead to increased heat stroke counts. Third and finally, it may be the case that the alerts induced changes in behavior that led to "real" increases in actual heat stroke. For example, it could be that on days when a warning was issued, various locations (e.g. workplaces, offices, schools) closed more, and perhaps individuals are better protected from heat stroke at these locations vs. at their own home (e.g. due to better air conditioning). More generally, it may be that other factors associated with the alerts encourage individuals to engage in an "adverse" behavioral response, thus exposing them to greater risk for heat stroke in response to an alert.

In order to shed some light onto these potential mechanisms, we first utilize micro-data from the population of ambulance records in Japan.[11] We first find that total ambulance calls increase in response to a heat alert, even for cases where the patient was "severely injured." Thus, the main effects cannot be solely attributed to health diagnoses "substituting" away from other sudden illnesses and into heat stroke (since substitution would imply no change in ambulance pickups). Additionally, we estimate a small but statistically insignificant increase in overall mortality (that involved an ambulance) in region-days with a heat alert. Combined with a statistically significant positive effect for severe cases (in both heat strokes and overall ambulance records), we further believe that the overall heat stroke effect cannot be solely attributed to "increased reporting" (since severe cases, and cases which lead to death, would presumably have utilized an ambulance irrespective of an alert being issued).

Thus, to consider the possibility of an "adverse" behavioral response to the alerts, we next make use of data from Google Trends and Google Community Mobility Reports, which tracked daily foot traffic across locations by several broad categories from 2020 to 2022, while also considering the home residence and workplace of the tracked individual. From Google Trends, we first find that the heat alert raised general awareness, as reflected by increases in searches for the terms "heat stroke," "heat stroke alert," and "temperature." Searches for "air conditioning" also significantly increased. Searches for the general term "indoor"

[11]Discussed in further detail later, and as described in Akesaka and Shigeoka (2023), ambulance services are provided to the public for free in Japan, and thus constitute a significant share of hospital transport in Japan.

increase as well, though searches for specific popular indoor activities such as “cinema” and “karaoke” were unaffected. These Google Search results may reflect some evidence of “avoidance” behaviors, the specific aim of the heat-health alert in Japan.

The remainder of the evidence, however, suggests a potentially adverse response to the heat alert. First, in the Google Mobility Report data, people were *more* likely to leave their home on alert days. This behavior stands in direct contrast to the government guideline to stay home and use air conditioner. Although this may simply reflect the “increased reporting” channel (where people leave their home via ambulance and report cases that were otherwise unidentified), we also estimate a large increase in visitations to parks, the sole outdoor category provided in the Google Community Mobility Reports, when an alert was issued. We also find increased searches on Google for outdoor activities such as “sea bathing” and “park.” It is also possible that these behaviors still constitute “avoidance” behaviors, where perhaps people are better protected from high heat outside their own homes. However, additional explorations fail to uncover any evidence of avoidance behaviors leading to improved outcomes. For example, we find that the alerts have no effect on heat strokes on days $t + 1$ or $t + 2$, where perhaps avoidance behaviors today would capitalize into reduced heat strokes tomorrow or the day after tomorrow.[12]

To further contextualize our results, we consider several related studies from the Japanese context which uncovered similarly concerning results. First, a recent study from He and Tanaka (2023) found increased mortality and reduced air conditioning usage, particularly on extremely hot days, in response to large-scale energy-saving campaigns in Japan. Related work from Neidell et al. (2021) found that households reduced their electricity consumption when local prices rose (after the Fukushima accident), causing an increase in mortality. These studies suggest that in our context, the observed behavioral responses may come from a lack of (good) air conditioning at home and/or a desire to save energy on (what households perceive to be) especially hot days. For example, individuals may have been more electricity-price conscious when a warning was issued. We also find that people are more likely to leave their home and more likely to go to the park in response to actual WBGT - these behaviors are consistent with households saving energy on high heat days, and thus households may be using the heat alert as an indicator for whether they should leave their home in order to save energy.[13]

This paper provides several important contributions to the literature. To our knowledge, it is the first large-scale study to examine the causal effects of a heat warning system. It is also one of only two studies

[12] We further find no change in the likelihood an individual went into work. We also note that the effect of a heat alert on heat stroke is slightly weaker on weekdays than on weekends, suggesting that changes in workplace location on heat alert days cannot explain the results.

[13] Other related work from Akesaka and Shigeoka (2023) also utilizes geolocation data in Japan to find that despite causing significant mortality risks, people do not curtail outdoor activities in response to seasonal allergies.

to examine the effects of *any* public alert system on direct health outcomes, with nearly the entirety of the previous literature focusing on how air quality alerts impact avoidance behavior.[14] Early work from Neidell (2009) and Zivin and Neidell (2009) examine avoidance behavior in California in response to smog alerts, finding reduced outdoor activity when alerts are issued, with reduced efficacy for alerts made on consecutive days. Noonan (2014) finds that elderly users and exercisers reduce their use of a major park in response to a smog alert, with no effect on driving behavior. Ward and Beatty (2016) find similar behavioral effects of reduced outdoor activity, particularly among the elderly. Saberian et al. (2017) find reductions in cycling in Sydney when an air quality alert is issued. Beyond air quality alerts, Gutteling et al. (2018) investigate behavioral responses from 643 survey participants to a Dutch civil defense warning system, carried via cell phone messages, which warned citizens of various local natural and man-made threats (e.g. toxic fumes released with fire). Ferris and Newburn (2017) find that mobile phone alerts for flash flood events in Virginia led to reductions in traffic volume and car accidents.[15]

Our study also relates to the strand of literature identifying factors that influence the climate-health relationship. As mentioned above, perhaps the two most directly related studies to ours come from Neidell et al. (2021) and He and Tanaka (2023), who study electricity prices and energy-saving campaigns, respectively, in the Japanese context.[16] Other prominent studies identifying adaptive behaviors to climate include Deschenes and Moretti (2009) (migration), Zivin and Neidell (2014) (staying indoors), and Barreca et al. (2016) (air conditioning).[17] Finally, Mullins and White (2020) show that better access to health care can mitigate the negative impacts of heat. Our study contributes to this literature by being the first to look at how public alerts affect the climate-health relationship, and along with He and Tanaka (2023), is one of the first studies to evaluate any government policy specifically targeting behavioral responses to higher temperatures.[18]

The remainder of this paper proceeds as follows. In Section 2, we describe Japan's heat warning system

[14]To our knowledge, the only study examining how a public alert impacted direct health outcomes comes from Mullins and Bharadwaj (2015), who find that an "Environmental Episode" policy from Chile reduced air pollution by 20% and improved day-of elderly mortality.

[15]A separate strand of related research investigates how private health alerts affect health outcomes (Kim et al., 2019; Iizuka et al., 2021), finding that alerts overall have little effect on behavior, but can have substantial positive effects for those who are at high risk (for diabetes).

[16]Our study also differs from these two in several noticeable ways. First, we study different treatment variables (heat-health warnings vs. energy-saving campaigns vs. electricity prices). Further, whereas the previous two studies utilize monthly-regional variation driven by the Fukushima accident, our data are more granular (daily and smaller spatial units) while our identifying variation utilize both space-time variation and variation in whether forecasted WBGT exceeded a specific threshold. Our data also include health outcomes beyond mortality, including severity of heat stroke. Lastly, given the recent timeframe of our study, we are able to utilize Google Mobility Reports data to measure behavioral outcomes such as staying indoors.

[17]Additional studies broadly discuss the role of adaption on the climate-health relationship, including Deschênes and Greenstone (2011), Geruso and Spears (2018), Mullins and White (2019), and Heutel et al. (2021).

[18]A relevant working paper from Shrader et al. (2023) estimates the value of accurate weather forecasts in the United States. The authors estimate that a 50% increase in forecast accuracy would save approximately 2,200 lives a year. Our work compliments Shrader et al. (2023) by highlighting "false" weather warnings due to inaccurate forecasts could lead to further health losses.

and the criteria a warning is based on. In Section 3, we describe the various data used and present descriptive statistics. In Section 4, we introduce our estimation strategy. In Section 5, we present our results. We first discuss graphical findings before turning to our main results. We further discuss potential mechanisms. In Section 6, we show robustness of our results to a series of considerations. Finally, in Section 7, we conclude.

2 Background

In 2020, Japan introduced a comprehensive heat warning system to raise awareness of heat-related illnesses and promote heat stroke prevention measures. In the first year of implementation, the program was tested in nine prefectures (Chiba, Gunma, Ibaraki, Kanagawa, Nagano, Saitama, Tochigi, Tokyo, and Yamanashi), and then the warning system was rolled out to all regions of Japan starting in 2021. Alerts are communicated via the Ministry of Environment's homepage, weather news broadcasting stations, radio, public institutions, social networking services (i.e. LINE), and smartphone applications. The way these alerts are communicated (e.g. through newspapers, TV, public billboards and loudspeakers, or TV screens in trains) makes it likely that most people receive this information.[19] The alerts encourage people to stay at home, drink a lot, take salt, avoid exercise, use air-conditioning, and report to others if they are in poor physical condition (i.e. children, elderly, and disabled). As shown in Figure 1, Google Searches for the term "heat stroke alert" in Japan began when the warning was introduced and were especially high in the summer, and as shown later econometrically, daily searches in a region increased significantly when a warning was issued.

Warnings are issued based on whether the region-specific forecasted heat index "Wet Bulb Globe Temperature" (WBGT) (Yaglou et al., 1957) exceeds a threshold of 33.0°C.[20] This index serves as a more comprehensive indicator of heat stress and its impact on the human body than temperature alone, because it also incorporates humidity, radiant heat, and wind speed, which in combination with temperature influence the risk of heat stroke. In order to calculate forecasted WBGT for a specific region, the Japan Meteorological Agency (JMA) collects information on temperature, humidity, amount of rain, wind strength, and solar radiation across 1,300 weather stations. All of these stations collect information on rain, while 840 stations collect further information about temperature, humidity, and wind strength/direction. Among all weather stations, only 11 stations collect information about wet bulb temperature, wbt, black bulb temperature, bbt,

[19] In the online appendix, we present examples of those public announcements.

[20] Website: https://www.env.go.jp/press/109467.html, retrieved October 10, 2022.

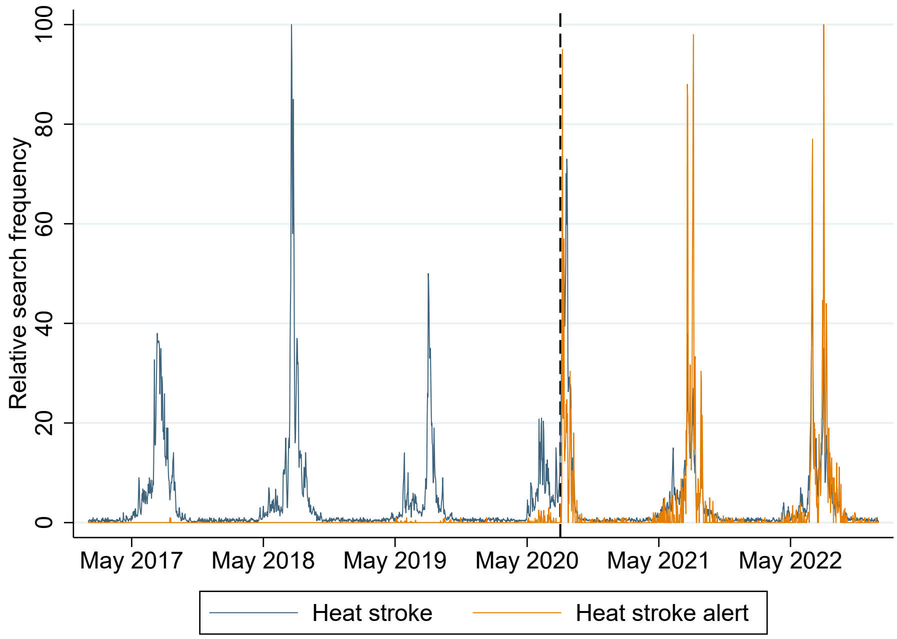


Figure 1: Google Searches for Heat Strokes and Heat Stroke Alerts

Source: Authors' presentation using data from Google Trends. *Note*: This graph shows Google Trends Searches for 熱中症 (heat stroke) and 熱中症警戒アラート (heat stroke alert), respectively. The dashed line indicates the introduction of the Heat Stroke Alert System in August 2020.

and dry bulb temperature, dbt, and calculate WBGT according to

$$\text{WBGT} = 0.7 \cdot wbt + 0.2 \cdot bbt + 0.1 \cdot dbt. \tag{1}$$

44 stations collect information about global solar radiation in order to estimate WBGT following Ono and Tonouchi (2014):

$$\begin{aligned} \text{WBGT} = {} & 0.735 \cdot Ta + 0.0374 \cdot RH + 0.00292 \cdot Ta \cdot RH + 7.619 \cdot SR \\ & - 4.557 \cdot SR^2 - 0.0572 \cdot WS - 4.064, \end{aligned} \tag{2}$$

with Ta being the temperature in $°C$, RH being the relative humidity in %, SR being global solar radiation in kW/m^2, and WS being the average wind speed in m/s. 98 stations do not collect information about global solar radiation, but estimate it using the relationship between the previous 10 minutes of sunshine hours (min) and the relationship between clear-sky irradiance and total solar irradiance based on historical observations. 687 stations further do not collect information about humidity. This information is estimated using numerical forecast data or predicted values from the JMA that is reanalyzed with observed values from surrounding observations. The remaining 460 stations only collect information about precipitation and therefore cannot estimate WBGT.

With a formula for WBGT for each of 840 weather stations, an alert is issued for a region based on whether *any* of the stations within a specific region have a forecasted WBGT greater than 33.0°C, i.e. whether the maximum forecasted WBGT across all stations within a region exceeds 33.0°C. Furthermore, the decision rule is based on two forecasts: 5pm of the previous day or 5am of the same day. For example, if the 5pm forecast for tomorrow's WBGT is 33.1°C for a single station in the Osaka region, then an alert is issued for Osaka (and will be broadcast across Osaka as a warning for tomorrow). If tomorrow, the same-day 5am forecast is at most 32.9°C across all stations in Osaka, then the alert still applies. In some cases, the 5pm forecast is below 33.0°C while the 5am forecast is above 33.0°C, in which case an alert is issued in the morning, even though no warning was issued the evening before. In our main analysis, for simplicity, we consider whether a warning was ever issued, regardless of the consistency between 5pm the previous day and 5am the same day, while later, for robustness, we separately consider the warnings issued based on the two forecasts.

3 Data and Summary Statistics

3.1 Heat Alerts

The Ministry of Environment provides information about the issuance of heat alerts for each of 58 regions on a given day since the implementation of the program in 2020.[21] As described in the previous section, these alerts are either issued at 5pm on the previous day or 5am on the same day. Appendix Figure A1 shows the fraction of dates between May 1 and September 30, 2022, that had a heat stroke warning by region. There exists substantial variation across regions in how often alerts were issued, with some regions issuing zero alerts (Aomori, Akita, Iwate, and Miyagi) and others issuing alerts for as many as 26% (Oita) of the summer days in 2022.

3.2 Wet Bulb Globe Temperature (WBGT)

Although heat alerts are issued on the regional level based on the maximum forecasted WBGT among all weather stations within a region, data from the Ministry of Environment only includes daily measurements for the maximum *realized* WBGT for each of 840 weather stations, rounded to the first decimal.[22] The

[21] Website: https://www.wbgt.env.go.jp/alert_record.php, retrieved October 10, 2022.
[22] Website: https://www.wbgt.env.go.jp/doc_trendcal.php, retrieved October 10, 2022.

data cover the dates between May 1 and September 30 for the years 2017 to 2022.[23] To reflect the decision variable, we calculate the maximum realized WBGT across all stations within a region on a specific date as our primary control variable. Appendix Figure A2 plots the full distribution of observed daily WBGT across region-days for the full sample.

Importantly, because we do not have information about the 5pm and 5am forecasted WBGTs, i.e. the "running variable" used to determine whether an alert was issued for a specific region, our identification strategy will effectively leverage variation in whether an alert was issued controlling for actual WBGT (described further in Section 4). In particular, alerts do not perfectly match actual WBGT because of forecasting errors. Thus, although we do not observe the actual decision variable for determining an alert, we are still able to flexibly control for the sole concerning confounder of the heat alert: actual WBGT.

3.3 Health Outcomes

3.3.1 Heat Stroke

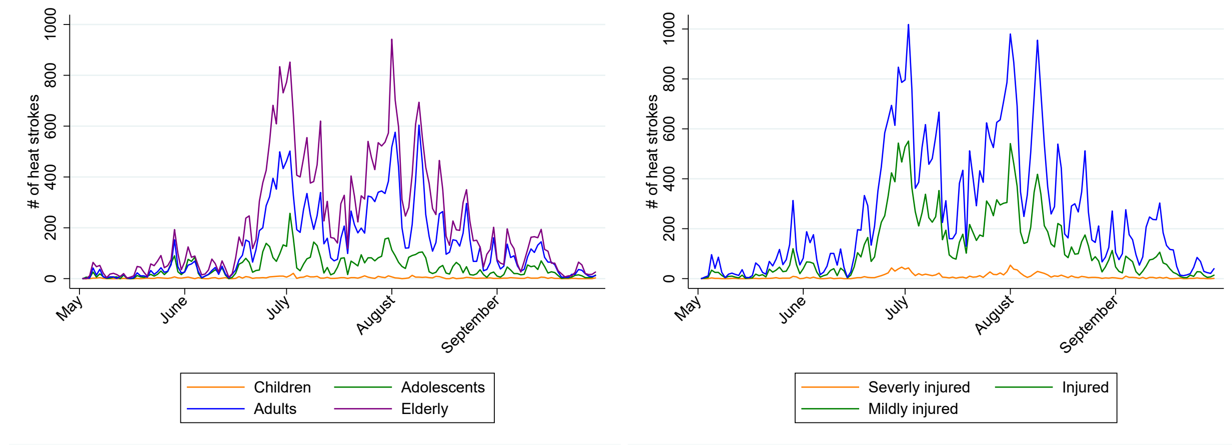
The primary outcome variable for our study is heat stroke. Heat stroke is the most serious heat-related illness (vs. heat exhaustion, rhabdomyolysis, heat syncope, heat cramps, and heat rash).[24] With a heat stroke, the body typically loses the ability to control its own temperature, which can lead to permanent disability or death without emergency treatment. Symptoms include very high body temperature, potential loss of consciousness, seizures, dry skin, or profuse sweating. Treatment typically includes cooled IV fluids, a cooling blanket, an ice bath, medication for seizures, or supplemental oxygen. Broadly speaking, there are two types of heat strokes: "exertional" heat stroke results from physical overexertion in hot, humid conditions, and develops over a few hours, while non-exertional, or "classic," heat stroke is the more common type, and is due to age or underlying health conditions, and tends to develop over several days.

The Fire and Disaster Management Agency (FDMA) provides administrative data on the number of heat strokes by age group, place of incidence, and severity, for each of 47 prefectures on a daily basis in the summer months since 2008.[25] These data include all individuals who were transported to a hospital via an ambulance. As Akesaka and Shigeoka (2023) describe, ambulance services are provided to the public for free, and thus, utilized similarly by individuals across socioeconomic status. Ambulance use is extremely prevalent among patients in Japan, and in general, Japanese policymakers express significant concern of

[23] Detailed data can be downloaded following the description in the manual: https://www.wbgt.env.go.jp/man15NH/R04_wbgt_data_service_manual.pdf, retrieved October 10, 2022.

[24] Source: Center for Disease Control and Prevention, 2023. Website: https://www.cdc.gov/niosh/topics/heatstress/heatrelillness.html, retrieved July 13, 2023.

[25] Website: https://www.fdma.go.jp/disaster/, retrieved October 10, 2022.



(a) Age (b) Severity

Figure 2: Heat Strokes by Age and Severity

Source: Authors' presentation using data from Fire and Disaster Management Agency (FDMA). *Note*: This graph shows the number of heat strokes by age group (a) and severity (b) in 2022 for Japan. "Mildly injured" are any heat stroke incidents that did not require hospitalization. "Injured" are incidents that required no more than three weeks of hospitalization, while "severely injured" are those that required at least three weeks of hospitalization.

overuse of ambulance services. For example, Kadooka et al. (2017) estimate that nearly 50% of ambulance rides were for minor conditions where patients could have taken a taxi instead of an ambulance. In response to the summer heat of 2023, Tokyo is on pace to set a new record number for ambulance dispatches.[26] Thus, though we do not observe visits made to hospitals via methods other than an ambulance, the data likely cover a significant share of all visits to hospitals.

In Figure 2, we plot counts of daily heat strokes by age group and severity for the entirety of Japan in 2022. Consistent with the prior health literature, elderly are at the most risk for heat stroke, and heat stroke incidence is strongest in July and August. Although most cases are mild, there are a few severely injured cases (which require multiple weeks of hospitalization).

We match these prefecture-daily data to the Ministry of Environment's region-daily data. Three prefectures (Hokkaido, Kagoshima, and Okinawa) are distinguished into 8, 2, and 4 regions, respectively, whereas the remaining 44 prefectures perfectly map into the remaining 44 regions. For simplicity, we drop these three prefectures. Since the Ministry of Environment's data date from 2017, our matched sample begins in 2017, granting us an additional three years of observations pre-warning system. In Appendix Figure A3, we overlap counts of heat strokes with heat alerts for all of Japan for 2021 and 2022. Unsurprisingly, a positive correlation arises since heat alerts are issued on hotter days, and heat stroke incidence increases on hotter

[26]Source: The Japan Times. Website: https://www.japantimes.co.jp/news/2023/07/01/national/tokyo-ambulance-calls-rise/, retrieved September 7, 2023.

days.

3.3.2 Ambulance Transports

The FDMA also provides administrative data on the population of ambulance transportations to hospitals in Japan. The data cover all transports from 2017 to 2021 for each of the 47 prefectures on a daily basis. Appendix Table A1 provides summary statistics for the number of transports by severity. There are on average 371 ambulance transports per day and region. The majority of these transports are for mild cases (92%), while 7% are for severe cases, and 1% result in death. Similar to the heat stroke severity categorization, "severe" cases require multiple nights of hospitalization, while "mild" cases require no hospitalization.

We utilize these data for multiple purposes. First, assuming the majority of cases that result in death in Japan utilize an ambulance, these records allow us to observe the effect of the heat alerts on region-day-level mortality. Second, discussed in further detail later, these data allow us to investigate potential substitution between heat stroke cases and other sudden illnesses. In particular, we can distinguish between injuries resulting from natural disasters, fire, or water (0.1%), traffic accidents (6.9%), work- or sports-related accidents (2.0%), sudden and general illnesses (81.7%), injuries due to violence and self-harm (1.0%), and transfers from one hospital to another (8.2%). Since heat stroke falls under the "sudden and general illnesses" category, the other categories also grant us useful placebo tests.

3.4 Behavioral Outcomes

3.4.1 Google Trends

We further utilize data from Google Trends to examine how searches for various terms on Google's search engine changed in response to heat alerts. Google Trends provides measurements of the relative search frequency of a queried search item (e.g. heat stroke alert) on Google Search for a selected geography (e.g. Tokyo) and time period (e.g. May 1 to September 30, 2021), indexed to a range of 0 and 100, where 100 represents the date within the specified geography that had the greatest search volume for the queried search item. Importantly, these indices are generated from relative search volume, not total searches on Google. Google Trends takes a random sample of total searches for the queried geography and time frame, then generates the indices based on the share of total searches that the queried item constituted. We consider searches for the terms heat strokes, heat stroke alert, weather, temperature, air conditioner, outdoor, sea bathing, park, indoor, cinema, and karaoke for each prefecture for May 1 to September 30 for 2020 to 2022, respectively. Appendix Table A2 provides summary statistics for these terms.[27]

3.4.2 Google Community Mobility Reports

In response to the COVID-19 pandemic, Google provided aggregated, anonymized "reports" which tracked daily movement by location. The data used to generate the reports are the same used in Google Maps, where users can track how busy a specific location is. In particular, the Community Mobility Reports include information on shares of people who left their residence, and which locations individuals visited across different categories: retail and recreation, groceries and pharmacies, parks, transit stations, and workplaces. These reports were provided from January 2020, with daily updates ceasing on October 15, 2022. Visitations are reported as percentage changes in daily visits relative to visitations for that weekday pre-Covid. Naturally, these data have been used to investigate questions related to traveling behaviors in response to COVID-19 and corresponding policies (e.g. Chetty et al., 2020; Fernández-Villaverde and Jones, 2020; Sulyok and Walker, 2020; Karaivanov et al., 2021; Mendolia et al., 2021). For our study, we utilize mobility reports at the region-day level in order to investigate behavioral responses to the heat alerts.[28] Since these data only start from 2020, analyses using the Community Mobility Reports do not include any pre-treatment observations.[29] Appendix Table A3 provides summary statistics for these reports.

3.5 Sample Restrictions

Our heat stroke data consist of 44 regions with information about maximum realized WBGT and heat alerts for each day between May 1st and September 30 for 2017 to 2022. However, because there are almost no heat alerts for region-days with a low WBGT, our primary analyses restrict the data to days with a WBGT of at least 28$°C$. Further, there are a few region-days with an extremely high number of heat strokes; to lessen the effect that potential outlier events drive our results, we drop region-days with heat strokes above the 99th percentile. Robustness exercises test the sensitivity of our main results to these two sample selections. Our final primary sample consists of 20,612 observations.

[27]We downloaded the data for each prefecture and year (May 1 to September 30) separately, such that numbers are relative to the searches within a specific prefecture in a year. Alternatively, one could download all regions on a given date and repeat this for all required dates. Since the data are indices related to a specific geographic and temporal area, the numbers can change. Nevertheless, our results are robust to downloading the data for all regions of Japan at different points in time (not reported).

[28]To our knowledge, our study is the first to use these data to investigate mobility responses outside of Covid-related policies.

[29]More information on these reports can be found at https://www.google.com/covid19/mobility/.

3.6 Summary Statistics

Table 1 provides summary statistics for the months of May through September for the years 2017 to 2022 at the region-day level. There are approximately 15 heat strokes every day in any given region. The majority of heat strokes result in mild injury, which are cases that do not require overnight hospitalization. Only a few cases require hospitalization of up to three weeks. While only very few heat strokes are experienced by children, over half of all heat strokes are experienced by those aged 65 years and older. Interestingly, the most common location for a heat stroke to occur is in one's home. This, coupled with the fact that elderly are at greater risk of heat stroke, highlights how the majority of heat strokes occur due to age and underlying health, and are not necessarily due to one overexerting themselves outdoors. Heat alerts were issued on approximately six percent of region-days. This number increases to almost 17% when only considering the years 2021 and 2022, the time the alert system was fully implemented. Finally, the average WBGT in the summer months was around $31°C$.

4 Empirical Strategy

A heat stroke warning was issued for region r on day t when the maximum forecasted WBGT from either the 5pm forecast on day $t-1$ or the 5am forecast on day t exceeded $33.0°C$ at any weather station within region r. Conversely, a region issued no warning if $\text{WBGT}_{rt} < 33$ for both the evening-before (5pm) and the morning-of (5am) forecasts across all stations within region r.

If one observed the forecasted WBGT, one could compare region-days with a forecasted WBGT just marginally below the $33.0°C$ threshold against region-days with a forecasted WBGT just marginally above the threshold in a Regression Discontinuity Design (RDD). This identification strategy would then compare region-days with very similar forecasts that differed within some tight range of forecasted WBGTs, and assume that region-days on opposite sides of the cutoff are otherwise similar to each other. However, as discussed in the previous section, our data do not include station-level forecasts of WBGT, and thus, we do not observe the "running variable" necessary for an RDD.[30]

Still, our data include realized WBGT and indicators for whether a heat stroke warning was issued due to the evening-before or morning-of forecasts, or both, i.e. whether the forecasted WBGT crossed the

[30] Moreover, the most concerning confounder for identifying the effect of an alert on heat stroke would be actual WBGT, not forecasted WBGT, since actual weather is what directly impacts individuals. Still, it is possible that individuals make decisions (e.g. plans for tomorrow) based on forecasted WBGT, independent of the alert, and fail to adjust their plans when actual WBGT differs from forecasted WBGT. So long as the forecasting errors are distributed around zero, any such possibility should merely attenuate estimated treatment effects of an alert on outcomes. Also note that forecasted WBGT should theoretically be perfectly collinear with the alert, and thus it is impossible to disentangle the effect of the alert from the effect of forecasted WBGT.

Table 1: Summary Statistics

	(1) Obs.	(2) Mean	(3) SD	(4) Min	(5) Max
Panel A: Outcome variables					
Heat strokes: total	20612	14.890	19.95	0.00	141.00
Heat stroke result: mild injury	20612	9.443	13.17	0.00	115.00
Heat stroke result: injury	20612	5.020	7.21	0.00	71.00
Heat stroke result: severe injury	20612	0.340	0.84	0.00	15.00
Heat strokes: 28 days–6 years	20612	0.112	0.38	0.00	6.00
Heat strokes: 7–17 years	20612	1.690	3.07	0.00	54.00
Heat strokes: 18–64 years	20612	5.196	7.89	0.00	79.00
Heat strokes: $\geq$65 years	20612	7.891	10.39	0.00	95.00
Heat stroke loc.: home	20612	5.999	8.26	0.00	86.00
Heat stroke loc.: workplace (construction)	20612	1.669	3.04	0.00	43.00
Heat stroke loc.: workplace (field)	20612	0.364	0.75	0.00	10.00
Heat stroke loc.: education	20612	0.825	1.83	0.00	42.00
Heat stroke loc.: public (inside)	20612	1.132	2.20	0.00	35.00
Heat stroke loc.: public (outside)	20612	1.803	3.08	0.00	56.00
Heat stroke loc.: street	20612	2.272	4.16	0.00	68.00
Heat stroke loc.: other	20612	0.826	1.37	0.00	16.00
Panel B: Treatment variable					
$=1$ if heat alert	20612	0.063	0.24	0.00	1.00
Panel C: Control variable					
WBGT index	20612	31.212	1.86	28.00	37.60

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table presents descriptive statistics of all key variables used in the analysis. Mild injuries do not require overnight hospitalization. Severe injuries require up to three weeks of inpatient care. Home: all locations on the property. Workplace (construction): road construction sites, factories, manufacturing, etc. Workplace (field): fields, forests, oceans, rivers, etc. (only if agricultural, livestock, or fishery operations are conducted). Education: kindergartens, nursery schools, elementary schools, junior high schools, high schools, vocational schools, universities, etc. Public (inside): indoor areas where people enter and exit (theaters, concert halls, restaurants, department stores, hospitals, public bathhouses, train stations (underground platforms), etc.). Public (outside): outdoor portions of places where people enter and exit (stadiums, outdoor parking lots, outdoor concert venues, train stations (outdoor platforms), etc.). Street: general roads, sidewalks, toll roads, highways, etc. WBGT ≥ 28.

threshold or not. With these data, we are able to regress our outcome variables y_{rt} on a dummy variable

$Alert_{rt}$, indicating a heat stroke warning in region r at time t, while flexibly controlling for actual WBGT:

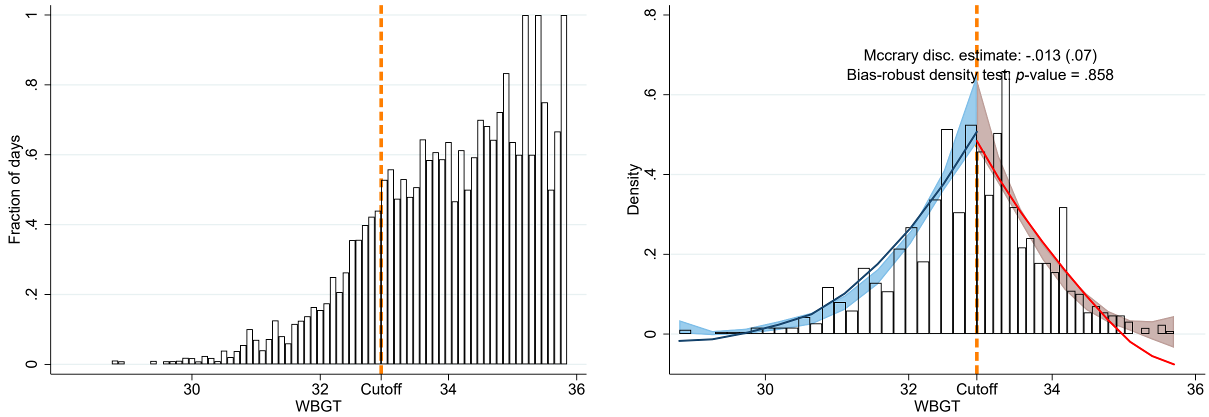
$$y_{rt} = \alpha + \beta Alert_{rt} + \sum_{k=28.0^{\circ}C}^{K} \gamma_k \mathbb{1}\left(\mathrm{WBGT}_{rt} = k\right) + \delta_r + \lambda_t + \zeta_r year + \epsilon_{rt}, \tag{3}$$

where $\mathbb{1}(\cdot)$ is an indicator function, δ_r are region fixed effects, λ_t are day fixed effects, and ζ_r are region-specific linear time trends. To control for WBGT, we follow the decision rule for the alerts and assign each region-day rt the maximum observed WBGT across stations within that region-day (calculated as WBGT_{rt}), and include dummies for each possible WBGT_{rt}, rounded to the nearest tenth.[31]

We are able to separately estimate β from γ_k since $Alert_{rt}$ does not perfectly correlate with WBGT_{rt}. In other words, we are able to separately disentangle the effects of the heat warning from heat itself on outcomes since (unobserved) forecasted WBGT_{rt} does not perfectly predict actual (observed) WBGT_{rt}. In essence, our identification strategy relies on variations in forecasted and actual WBGT_{rt} (i.e. "forecasting errors"), and whether forecasted WBGT_{rt} and actual WBGT_{rt} laid on the opposite sides of the 33.0$^{\circ}C$ threshold. Because of these forecast errors, we observe region-days where an alert was issued but actual WBGT_{rt} was below 33.0$^{\circ}C$ (i.e. false positive), and region-days where no alert was issued but actual WBGT_{rt} exceeded 33.0$^{\circ}C$ (i.e. false negative). This identification strategy relies on the assumption that "forecasting errors" are random and therefore independent from unobservable features of region-days that could potentially affect heat stroke. Region fixed effects further control for all time-invariant differences across regions, such as average temperature and average heat stroke incidence. Day fixed effects control for all unobserved nationwide trends, including potential changes in average weather or heat stroke incidence. Since the heat stroke warnings are only based on forecasted WBGT_{rt} and no other potentially confounding factors, controlling for actual WBGT_{rt} in this panel regression specification allows us to identify the causal effect of the heat warning on our outcomes y_{rt}.

Our identification strategy relies on the fact that regions issue alerts based on the forecasted WBGT across all stations, yet due to forecasting errors, the same region may experience different days with identical WBGTs (e.g. 33.1$^{\circ}C$) but some with a warning issued and others without. To illustrate the relationship between WGBT and alerts, in panel (a) of Figure 3, we plot the fraction of region-days that issued an alert by the WBGT for that region-day. We see that around the 33.0$^{\circ}C$ cutoff, the probability an alert is issued increases from around 45% to around 55%, and slowly increases in WBGT. Still, due to forecasting errors, substantial variation remains: some region-days with maximum WBGT as low as 30.0$^{\circ}C$ issued alerts, while some region-days with maximum WBGT above 34.0$^{\circ}C$ failed to issue an alert. If forecasting errors

[31] Unreported in equation (3) for simplicity, we similarly flexibly control for WBGT_{rt-1}, the region's previous day's WBGT.



(a) Fraction of region-days with alert

(b) Distribution of false alerts

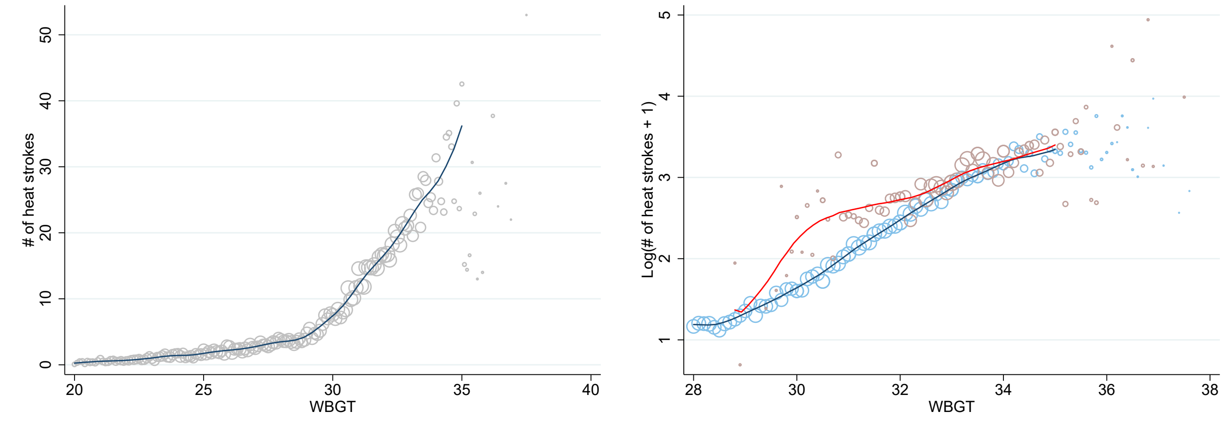
Figure 3: Alerts and WBGT

Source: Authors' presentation using data from Ministry of Environment. *Note*: Figure (a) shows the share of region-days with an alert among all region-days for all values of WBGT. Figure (b) shows the distribution of false alerts for different values of WBGT. Local polynomial approximations on both sides of the cutoff are based on the triangular kernel and local quadratic regressions. Bandwidths for the local polynomial regressions on both sides are 2.35. 95%-confidence intervals are shown as shaded area. $N = 1125$. Data for 2020–2022. In 2020, the sample is restricted to the nine regions that had the heat warning system implemented. WBGT $= [28, 36]$. Results are obtained from the Stata command `rddensity` of Cattaneo et al. (2018).

did not exist, then the entire mass of observations would lie above 33.0$°C$.

The core assumption is that regions do not systematically issue "false" alerts with respect to the 33.0$°C$ threshold. More specifically, we assume that regions are not more or less likely to issue an alert when WBGT_{rt} is below 33.0$°C$ relative to failing to issue an alert when WBGT_{rt} exceeds 33.0$°C$. This is equivalent to assuming that forecasting errors are random. In panel (b) of Figure 3, we present the distribution of false alerts across actual WBGT.[32] If forecasting errors are independent from actual WBGT, the distribution of false alerts should be smooth around the threshold. Conversely, if regions manipulate the forecasts or the decision rule such that more alerts are issued on either side of the 33.0$°C$ threshold, then we would expect to see a higher share of false alerts just above or below the threshold. To formally test for such a discontinuous jump in the distribution, we conduct the manipulation test of Cattaneo et al. (2020) and fail to reject the null hypothesis of no manipulation. Moreover, unsurprisingly, there is a greater mass of false alerts around the 33.0$°C$ cutoff, since this is where even a small forecasting error can result in a false alert. Thus, though we do not observe forecasted WBGT, this evidence suggests that is highly unlikely that forecasting errors are systematically biased away from zero, or that regions implementing the alerts systematically deviated from the 33.0$°C$ decision rule.

[32] A region-day issued a "false" alert if either (a) an alert was issued but actual WBGT was below 33.0$°C$, or (b) an alert was not issued but actual WBGT was greater than 33.0$°C$.



(a) WBGT and heat stroke

(b) WBGT and heat stroke by heat alert days

Figure 4: Relationship Between Heat Stroke and Wet Bulb Globe Temperature

Source: Authors' presentation using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: These graphs show the relationship between the WBGT index and heat stroke. Panel (b) separates the relationship by region-days with a heat alert (in red) vs. without a heat alert (in blue). Local linear regressions are based on a triangular kernel and the rule-of-thumb bandwidth.

5 Results

5.1 Descriptive Figures – Increased Heat Strokes on Alert Days Conditional on WBGT

We begin by investigating the graphical relationship between temperature, heat strokes, and heat alerts. In panel (a) of Figure 4, we plot the relationship between WBGT and the number of heat strokes across the full bandwidth of possible WBGTs. The size of the circles indicate the number of region-days that experienced that specific WBGT; for example, WBGTs above 35°C were relatively rare. Unsurprisingly, the relationship between WBGT and heat stroke is positive, and interestingly, appears to be nonlinear, where the number of heat strokes increases at an increasing rate for higher WBGT.

Then, in panel (b), we plot the relationship between WBGT and heat stroke separately for region-days without an alert (in blue) and with an alert (in red). Immediately, even in just the "raw" data, we see a preview of the main results: region-days when a heat alert was issued systematically have *more* heat stroke hospitalizations across the full range of WBGTs. Although this difference is largest for smaller values of WBGT, it remains positive for higher values with more region-days experiencing an alert. In the next section, we will test whether this descriptive result still manifests in our econometric models.

Table 2: Effect of Heat Alerts on Heat Strokes

	Log (# of heat strokes +1)		# of heat strokes			
	OLS				Neg. binomial regression	
	(1)	(2)	(3)	(4)	(5)	(6)
= 1 if heat alert	0.199***	0.197***	3.876***	4.108***	0.174***	0.172***
	(0.018)	(0.018)	(0.977)	(1.021)	(0.016)	(0.015)
WBGT	✓	✓	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓
Region-spec. trend		✓		✓		✓
Mean heat strokes	14.100	14.100	14.100	14.100	14.100	14.100
Observations	20612	20612	20612	20612	20612	20612

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from OLS and negative binomial regressions of the number of heat strokes in a region on the presence of a heat alert and a set of control variables. The mean number of heat strokes is for regions without heat alerts (control group). WBGT $\geq$ 28. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

5.2 Main Econometric Results

We begin by estimating Equation (3) for our full sample in Table 2, where we regress the incidence of heat strokes across region-days in response to a heat alert being issued, also controlling nonparametrically for the maximum region-day WBGT and region-previous-day WBGT. Across columns, we consider combinations of measurements of the outcome variable, regression models, and the sensitivity of the results to region-specific time trends. We cluster standard errors at the region level. Starting with the first two columns, which take the log of region-day heat strokes (plus one), we estimate that the incidence of heat strokes increased by nearly 20% for region-days that issued a heat alert. In columns (3) and (4), we estimate the count of heat strokes as the outcome variable to find similarly large effects. Since our primary outcome variable is count data (and includes zero), perhaps the most appropriate regression model is the negative binomial regression model, which we consider in columns (5) and (6).[33] For our fully specified model in column (6), the expected log count of heat strokes on heat alert days is 0.172 higher than the expected log count for region-days without an alert. Across all specifications, our estimates are statistically significant at the 1%-level. In the remainder of our analyses, we will estimate our full model using negative binomial regressions.

[33] Results do not change if we use a Poisson regression model instead.

To present the heat-health relationship, in Appendix Figure A4, we plot the coefficients from the same-day and lagged WBGT dummy variables. Unsurprisingly, higher WBGTs produce more heat strokes (conditional on the previous day's WBGT and on whether a heat alert was issued). We also see some evidence that lagged WBGT impacts heat stroke, which is consistent with "classical" heat stroke, where the consequences of heat exposure accumulate over several days. Still, the same-day impact of WBGT on heat stroke is substantially stronger.

5.3 Potential Mechanisms

In this section, we consider several exercises in order to unpack the potential mechanisms of our main results. Before moving to the analyses, we first hypothesize several channels of potential responses resulting from the heat warning. First, as is the primary intention of the program, we may expect some type of "avoidance" behavior - in response to an alert, individuals engage in behaviors to avoid or mitigate the consequences of exposure to high heat. Generally speaking, this would involve the movement of individuals from hotter to cooler areas, or greater usage of appliances to reduce ones temperature (e.g. air conditioning). Of course, this channel cannot explain our main results in Table 2, since such avoidance behavior should mitigate exposure to heat and thus reduce heat stroke. Still, avoidance behavior could exist, yet be outweighed by other opposing factors.

Thus, we further posit a secondary set of behavioral responses that may emerge due to a heat alert and could plausibly explain our findings. The first we call an "increased reporting" effect. Given the alert is explicitly about heat stroke health, it may be that the alert raises one's awareness of the potential presence of heat stroke. That is, in the absence of an alert, some people who experience a heat stroke may not realize it or may fail to consider the seriousness of their symptoms - subsequently, their heat stroke case goes unreported and uncounted. This would be consistent with a health literature that suggests that the damages from high heat are significantly under-counted and unreported (Ebi et al., 2021). Then, in the presence of the alert, some share of these individuals may now visit the hospital and report their heat stroke case.

The second channel we hypothesize is a "substitution" effect in how health cases are diagnosed. Heat stroke is generally brought upon by high heat, while high heat also substantially increases the risk of other sudden respiratory and cardiovascular diseases (e.g. Michelozzi et al., 2009). In the US, the EPA estimates that approximately one quarter of heat-related deaths are tied to cardiovascular disease interacting with high heat.[34] Given these important interactions, we posit that heat stroke diagnoses can potentially be

[34] Source: United States Environmental Protection Agency, https://www.epa.gov/climate-indicators/climate-change-indicators-heat-related-deaths, retrieved October 16, 2023.

conflated with other sudden illnesses when patients are transported to a hospital. Then, in the presence of the heat alert, the identification of heat-related symptoms by doctors and hospitals lead to "substitutions" away from other related sudden illnesses and into heat stroke. Furthermore, because doctors fill out a form providing information about the ambulance transport (i.e. cause or reason for the transport, severity, location of incidence, ...) upon arrival at the hospital, these on-the-spot decisions could also lead to substution away from other categories to sudden illnesses.[35]

Note that with both of these channels ("increased reporting" and "substitution"), the *actual* incidence of heat stroke is unchanged (all else equal), while *reported* heat stroke increases. Thus, it is possible that even in the presence of avoidance behavior (which subsequently reduces actual heat stroke), the total count of reported heat strokes increases in response to the alert (if the "increased reporting" and "substitution" effects outweigh reductions in actual heat stroke from any "avoidance" behaviors).

Lastly, we consider the possibility of an "adverse" behavioral response - the heat alert may be associated with other changing factors which then induce individuals to engage in behaviors that place them at a greater risk of heat stroke. For example, it may be that on days when a warning was issued, workplaces closed more, and perhaps individuals are better protected from high heat at their workplace relative to their home (perhaps due to better air conditioning).[36] If people were less likely to go into work, then they may have engaged in other adverse behaviors with their newfound free time as well. Moreover, He and Tanaka (2023) found increased mortality in high heat months after the Fukushima accident, which they attribute to a behavioral response from households saving energy (where electricity prices increased). In our context, it may be that households associate energy savings and electricity prices with the heat alert, and thus perhaps individuals engage in adverse behaviors in order to save energy on (what they perceive to be) a high heat day. In general, forecasting errors in our setting - when actual WBGT does not coincide with forecasted WBGT (and thus, the heat warning) - may lead people to engage in behaviors placing them at a greater risk of heat stroke (and as highlighted by Shrader et al. (2023), inaccurate forecasts contribute to a significant share of annual deaths in the United States).

In an effort to identify these potential mechanisms, our next analyses consider (a) various heterogeneities in the heat stroke results, (b) intertemporal models of alerts on future heat strokes, (c) other health and mortality outcomes using the population of ambulance records, (d) searches on Google Trends, and (e) foot

[35] Anecdotal evidence shows that these decisions may also reflect the attitudes of doctors. For example, in the case of work-related accidents, managers and companies are held responsible for the victim's compensation, while in the case of a sudden illness, the medical expenses need to be covered by the victim itself.

[36] Other interactive effects could occur too: for example, an alert could induce people to go into work to receive the air conditioning ("avoidance" behavior), but then perhaps an increased exposure to peers increases the likelihood that one reports another as suffering from a heat stroke ("increased reporting").

Table 3: Effect of Heat Alerts on Heat Strokes by Severity

	# of heat strokes		
	(1) Mild injury	(2) Injury	(3) Severe injury
= 1 if heat alert	0.167***	0.161***	0.166***
	(0.016)	(0.023)	(0.059)
WBGT	✓	✓	✓
Lagged WBGT	✓	✓	✓
Region FE	✓	✓	✓
Day FE	✓	✓	✓
Region-spec. trend	✓	✓	✓
Mean heat strokes	8.984	4.723	.310
Observations	20612	20612	20612

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from negative binomial regressions of the number of heat strokes by severity in a region on the presence of a heat alert and a set of control variables. The mean number of heat strokes is for regions without heat alerts (control group). WBGT $\geq$ 28. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

traffic outcomes as measured through Google Community Mobility Reports.

5.3.1 Heterogeneity in Heat Strokes

In Table 3, we consider our full specification while splitting heat stroke counts by the severity of the heat stroke. We estimate positive coefficients across the spectrum of severity. Strong positive effects are found for mild injuries that do not require hospitalizations and injuries that require some hospitalization. Interestingly, we precisely estimate a large effect for "severe injury" as well, which requires at least three weeks of hospitalization. Assuming that severe heat stroke would occur irrespective of awareness of an alert, this suggests that the main effects cannot be solely attributed to an "increased reporting" effect.

In Table 4, we estimate our full specification while splitting the counts of heat strokes across four age groups: children (aged less than seven years), adolescents (aged seven to 17), adults (aged 18 to 64), and elderly (aged 65 and above). We observe large and statistically significant effects for all age groups except for children. Although adolescents are the most resilient to heat, an alert increases heat strokes for this age group as well. However, the largest increase in heat stroke incidence is experienced by adults and the elderly in particular. In Japan, workers are strongly encouraged to retire by the age of 60, and in the majority of settings, are forced to retire by the age of 65. Given these strong results for elderly, this suggests that the overall increase in heat stroke incidence cannot be solely attributed to changes in behavior involving the

Table 4: Effect of Heat Alerts on Heat Strokes by Age Group

	# of heat strokes			
	(1) Child	(2) Adolesc.	(3) Adult	(4) Elderly
= 1 if heat alert	0.001	0.121***	0.166***	0.181***
	(0.097)	(0.033)	(0.020)	(0.016)
WBGT	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓
Region FE	✓	✓	✓	✓
Day FE	✓	✓	✓	✓
Region-spec. trend	✓	✓	✓	✓
Mean heat strokes	.109	1.649	4.908	7.431
Observations	20612	20612	20612	20612

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from negative binomial regressions of the number of heat strokes for different age groups in a region on the presence of a heat alert and a set of control variables. The mean number of heat strokes is for regions without heat alerts (control group). WBGT $\geq$ 28. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

workplace.[37]

In Appendix Table A4, we consider heterogeneity by the location of the heat stroke incident. We find that the positive effects are consistent across all locations, both indoors and outdoors. Though one may expect stronger effects in outdoor locations, these location outcomes do not trace movement (e.g. going from outdoors to indoors) and are net of any potential avoidance or adverse behavioral effects (i.e. people with varying degrees of risk of heat stroke changing where they spend their day in response to the alert).

5.3.2 Intertemporal Effects of Heat Alerts

Next, we consider several models to test for potential intertemporal effects of the heat alerts. First, given heat strokes can develop over several days, it may be that any gains from avoidance behaviors are not realized until several days in the future. That is, if an alert encouraged avoidance behavior on day t, then perhaps we would see a reduction in heat strokes on day $t+1$ or $t+2$. In these same models, we also consider whether future heat alerts (on day $t+1$ and $t+2$) had an effect on heat strokes on day t. This largely serves as a placebo exercise, though given an alert for $t+1$ could be generated on day t based on the 5pm forecast, it is possible that an alert made for $t+1$ could lead to an "increased reporting" effect on day t. The results from this exercise are presented in Table 5. By and large, we find no evidence of an

[37] In the online appendix, we present various robustness checks for levels of severity and age groups.

Table 5: Intertemporal Effect of Heat Alerts on Heat Strokes

	# of heat strokes			
	(1)	(2)	(3)	(4)
$= 1$ if heat alert$_{t-2}$			-0.007	-0.007
			(0.022)	(0.021)
$= 1$ if heat alert$_{t-1}$	0.014	0.014	0.017	0.018
	(0.016)	(0.015)	(0.015)	(0.015)
$= 1$ if heat alert$_{t}$	0.163***	0.162***	0.159***	0.159***
	(0.015)	(0.014)	(0.015)	(0.015)
$= 1$ if heat alert$_{t+1}$	0.011	0.011	0.003	0.004
	(0.017)	(0.017)	(0.017)	(0.017)
$= 1$ if heat alert$_{t+2}$			0.028*	0.029*
			(0.016)	(0.016)
WBGT$_{t-2}$			✓	✓
WBGT$_{t-1}$, WBGT$_{t}$, WBGT$_{t+1}$	✓	✓	✓	✓
WBGT$_{t+2}$			✓	✓
Region FE	✓	✓	✓	✓
Day FE	✓	✓	✓	✓
Region-spec. trend		✓		✓
Mean heat strokes	14.109	14.109	14.109	14.109
Observations	20600	20600	20600	20600

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from negative binomial regressions of the number of heat strokes in a region on the presence of a heat alert and a set of control variables. The mean number of heat strokes is for regions without heat alerts (control group). WBGT ≥ 28. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

intertemporal effect of alerts on heat stroke: both past and future heat alerts have almost no effect on heat alerts on day t, with the entirety of the main effects still occurring in response to an alert on day t. These results suggest that any increases in avoidance behavior did not capitalize into a reduction of heat strokes in the future.

We next consider whether there were any dynamic effects based on the issuance of alerts across consecutive days. Zivin and Neidell (2009) evaluate avoidance behavior in response to smog alerts, and find that there was little effect of a smog alert being issued on the second day of consecutive smog alert days. In Appendix Table A5, in column (1) we first consider our baseline model with a regressor for an alert on days t and $t-1$ separately. In column (2), we then interact the dummies for the two alert days to find no additional interactive effect for alerts being issued on consecutive days. For completeness, the remaining columns consider further interactions for consecutive-day alerts, including two day gaps, and still we find no evidence of intertemporal effects.

Table 6: Effect of Heat Alerts on Ambulance Transports

	# of heat strokes							
	All		Mild		Severe		Deaths	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$=1$ if heat alert	0.015***	0.014***	0.015***	0.013***	0.019*	0.023**	0.005	0.015
	(0.004)	(0.004)	(0.004)	(0.005)	(0.011)	(0.012)	(0.033)	(0.030)
WBGT	✓	✓	✓	✓	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓	✓	✓
Region-spec. trend		✓		✓		✓		✓
Mean # of transports	368.584	368.584	338.898	338.898	26.103	26.103	3.584	3.584
Observations	16879	16879	16879	16879	16879	16879	16879	16879

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from negative binomial regressions of the number of ambulance transports by severity in a region on the presence of a heat alert and a set of control variables. The mean number of transports is for regions without heat alerts (control group). WBGT $\geq$ 28. Data for 2017–2021. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

5.3.3 Ambulance Records – Mortality and Substitution

We next turn to the ambulance transports data, which include information on the population of ambulance transportations made in Japan from 2017 to 2021. In Table 6 we estimate our Equation (3) for all ambulance transportations using negative binomial regressions (in columns (1) and (2)), then split by severity in the remaining columns: mild (columns (3) and (4)), severe (columns (5) and (6)), and death (columns (7) and (8)). Similar to our main results, we estimate that heat alerts led to increases in ambulance usage, primarily due to mild and severe cases. These results suggest that the main effects cannot be driven solely by a "substitution" effect, where the increase in heat strokes comes from a reduction in ambulance usage stemming from other sudden illnesses.

In the final two columns of Table 6, we find that the heat alert did not have a statistically significant effect on mortality. Again this suggests that if there were any avoidance behaviors, the benefits were not capitalized into reduced mortality. They also show that the increase in heat strokes did not result in a statistically significant increase in mortality - concern remains however on the size of the effect, which is positive and comparable in magnitude to the effect sizes for mild and severe cases. We conclude that at a minimum, this is no evidence that the heat alerts led to reductions in mortality.

We can further utilize the ambulance records to identify the effect of heat alerts on patient outcomes

(a) Percentage change

(b) Absolute change

Figure 5: Effect of Heat Alerts on Ambulance Transports

Source: Authors' presentation using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: Panel (a) shows results from negative binomial regressions of the number of ambulance transports by severity in a region on the presence of a heat alert and a set of control variables. Panel (b) shows these results multiplied by the daily mean number of transports for regions without heat alerts (control group). WBGT $\geq$ 28. Data for 2017–2021. 95%-confidence intervals are based on robust standard errors clustered at the regional level.

by the "cause" or reason for the ambulance transport. We consider six categories: hospital transfer (an ambulance transferring a patient from one hospital to another), violence and self-harm, sudden and general illness (which includes heat strokes), work and sports, traffic accidents, and fire and water and natural disasters.[38] In Figure 5, we plot the coefficients from estimating our full specification for each of these categories, while also considering subgroups by severity (mild, severe, death). In panel (a), we plot the coefficients and their 95%-confidence intervals directly from our main equation, while panel (b) scales the coefficients by the daily mean number of ambulance transports. Reassuringly, we find that the positive effect of heat alerts on ambulance usage is driven solely by sudden and general illnesses, whereas the remaining five categories (which effectively act as placebo tests) all display statistically insignificant effects. When split by the three categories of severity, only one of the 15 placebo estimates (mortality due to fire and water and natural disasters) displays statistical significance, which can likely be attributed to general randomness. This implies that there is no substitution within the sudden illnesses category that could explain the results. We also do not find reductions in other accidents and illnesses that could be substituted for sudden illnesses.

[38]The raw FDMA data provide finer categories than these six groupings. For example, sudden illness is separately reported from general illness, and accidents due to fire, water, and natural disasters are separated from each other. We group sudden illnesses with general illnesses since heat strokes could fall into either of these categories, and in these data we otherwise cannot distinguish whether a specific ambulance transport was directly related to a heat stroke. We group fire, water, and natural disasters together due to the nature of their causes and to increase their counts.

Table 7: Google Trends – Effect of Heat Alerts on Searches Related to Awareness

	(1) "Heat stroke"	(2) "Heat stroke alert"	(3) "Weather"	(4) "Temperature"	(5) "Air conditioner"
= 1 if heat alert	5.248***	3.182***	−5.548***	4.575***	2.925***
	(1.182)	(0.976)	(0.638)	(0.724)	(0.711)
WBGT	✓	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓
Region-spec. trend	✓	✓	✓	✓	✓
Mean outcome	18.538	8.104	52.966	21.114	34.853
Observations	10647	10647	10647	10647	10647

Source: Authors' calculations using data from Google Trends and Ministry of Environment. *Note*: This table shows results from OLS regressions of Google Trends Searches for 熱中症 (heat stroke), 熱中症警戒アラート (heat stroke alert), 天気 (weather), 気温 (atmospheric temperature), and エアコン (air conditioner), respectively, in a region on the presence of a heat alert and a set of control variables. The mean outcome is for regions without heat alerts (control group). WBGT $\geq$ 28. Data for 2020-–2022. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

5.3.4 Google Trends

We now turn to Google Trends to see whether region-day level searches for specific terms responded to an alert. In Table 7, using the 2020 to 2022 sample, we consider searches for the terms (Japanese equivalent) "heat stroke," "heat stroke alert," "weather," "temperature," and "air conditioner." We find large increases in searches for "heat stroke" and "heat stroke alert," highlighting people's awareness of the warning and for the primary outcome of interest. Interestingly, we see a significant increase in searches for the word "temperature" and a significant decrease in searches for "weather," suggesting perhaps a substitution away from a generic term ("weather") in favor of becoming more temperature-conscious on days when an alert was issued.[39] Further, we observe a strong increase in searches for the term "air conditioner," highlighting again people's awareness of heat on a day with an alert.

Next, in Table 8, we consider whether searches for activities related to potential avoidance behavior responded to a heat alert, since searches could be indicative of intentions for that day's activities (e.g. a reduction in searches for outdoor activities could indicate intentions to stay at home more). We first see an *increase* in searches for outdoor activities such as "sea bathing" and "park" in response to an alert being issued, suggesting perhaps people were more likely to go outside on (what they perceive to be) a high heat

[39] Anecdotal evidence also suggests that it is common for people to search for "weather" when they are interested in precipitation. This could further explain the strong decrease in search frequency if an alert has been issued and rain is unlikely.

Table 8: Google Trends – Effect of Heat Alerts on Searches Related to Avoidance Behavior

	Outdoor			Indoor		
	(1) "Outdoor"	(2) "Sea bathing"	(3) "Park"	(4) "Indoor"	(5) "Cinema"	(6) "Karaoke"
= 1 if heat alert	0.553	2.084*	1.208**	1.577*	0.015	−0.563
	(0.801)	(1.127)	(0.465)	(0.934)	(0.920)	(0.773)
WBGT	✓	✓	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓
Region-spec. trend	✓	✓	✓	✓	✓	✓
Mean outcome	22.090	11.583	30.559	13.779	28.653	39.573
Observations	10647	10647	10647	10647	10647	10647

Source: Authors' calculations using data from Google Trends and Ministry of Environment. *Note*: This table shows results from OLS regressions of Google Trends Searches for 屋外 (outdoor), 海水浴 (sea bathing), 公園 (park), 屋内 (indoor), 映画館 (cinema), and カラオケ (karaoke), respectively, in a region on the presence of a heat alert and a set of control variables. The mean outcome is for regions without heat alerts (control group). WBGT $\geq$ 28. Data for 2020-–2022. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

day. Still, these results should be interpreted cautiously since an alert could just be reflective of a change in the composition of Google Search users for that region-day, and searches do not necessarily reflect actual actions (e.g. an alert piques an individual's curiosity as to whether a nearby park is closed for the day). These actions may also constitute "avoidance behavior" in that options such as "sea bathing" and "park" may better protect some individuals from heat relative to their own homes. Finally, we estimate a positive but statistically insignificant coefficient for the search term "outdoor" in response to a heat alert.

Similarly, searches for the general term "indoor" increase, which suggests that people search for ways to combat heat with indoor activities. In contrast, searches for popular indoor activities such as "cinema" or "karaoke," which provide protection from heat due to air conditioning, seem unaffected by alerts. People are not more or less interested in those indoor activities if an alert was issued. However, we should be cautious again because these results only reflect search behavior and not actual visits.[40]

5.3.5 Google Community Mobility Reports

Finally, we turn to the Google Community Mobility Reports data to investigate whether foot traffic responded to heat alerts. The results are presented in Table 9. Recall that the outcome is measured as

[40] In Appendix Figure A5, we plot the coefficients on the dummies for maximum daily WBGT for each Google Search term. The relationship is positive for heat stroke, temperature, and air conditioner, suggesting that these searches are responsive to heat itself.

Table 9: Google Mobility Reports – Effect of Heat Alerts on Mobility

	Percentage change in visits					
	(1) Retail & recreation	(2) Grocery & pharmacy	(3) Parks	(4) Transit stations	(5) Workplaces	(6) Residential
$= 1$ if heat alert	0.083	0.323**	1.713***	0.945**	−0.029	−0.115**
	(0.224)	(0.146)	(0.613)	(0.400)	(0.142)	(0.056)
WBGT	✓	✓	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓
Region-spec. trend	✓	✓	✓	✓	✓	✓
Mean outcome	−6.729	4.511	2.640	−15.944	−11.262	5.780
Observations	10647	10647	10293	10636	10647	10647

Source: Authors' calculations using data from Google COVID-19 Community Mobility Reports and Ministry of Environment. *Note*: This table shows results from OLS regressions of the percentage changes in number of visits to specific places in a region on the presence of a heat alert and a set of control variables. The mean number of the outcomes is for regions without heat alerts (control group). WBGT $\geq$ 28. Data for 2020-–2022. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

the percentage change in visits to a specific place (e.g. "Retail & recreation") relative to pre-Covid;[41] for example, looking at the means of the outcome variables, people spent more time at home and at parks and less time at transit stations and workplaces relative to pre-Covid.

We find that people spent *less* time at home in response to an alert. This may be reflective of people visiting the hospital for heat stroke (i.e. "increased reporting"). But, we also find that people spent significantly more time at parks, which aligns with the Google Trends results from Table 8. This could be indicative of avoidance behavior, or it may be that people engaged in more (outdoor) activities, placing themselves at greater risk for heat stroke. In Appendix Figure A6, we plot the coefficients from the dummies for daily WBGT for each location category - interestingly, we also find that people are more likely to leave their home residence and more likely to go to the park in response to a higher WBGT. Importantly, we do not see any change in visits to workplaces in response to a heat alert. Thus, it does not seem to be the case that people are skipping their workplace more in response to an alert in favor of, for instance, going to the park.[42]

[41] In particular, the reports show percentage changes relative to a reference day, which is the median of the corresponding weekday in the five weeks preceding the COVID-19 pandemic (i.e. January 3 to February 6).

[42] In line with these null workplace results, in Appendix Table A6, we separate the main effects by weekdays and weekends and only find a slightly smaller effect of heat alert on weekdays.

5.4 Overall Mortality

In Appendix Figure A7, we provide further analyses utilizing additional data from the Vital Statistics reported by the Ministry of Health, Labor, and Welfare (MHLW) on the population of deaths in Japan, which are only available at the region-month level by cause, gender, and age at death.[43] In contrast to the mortality data from the ambulance records, these data include all deaths across Japan independent of ambulance usage. Since the data do not allow us to estimate our preferable model at the region-day level, we closely follow He and Tanaka (2023) to estimate monthly-level effects of heat alerts on mortality. In short, we find some evidence of increased mortality that is mostly driven by the elderly, particularly due to "external causes" and the "circulatory system," in months that had more heat alerts (controlling for monthly WBGT). Although death by heat stroke does not have a separate category, it is most likely categorized as a problem with the respiratory or circulatory system. This evidence once again suggests there was little avoidance behavior responses, and if anything, potentially adverse responses among the elderly.

6 Robustness Checks and Placebo Exercises

In this section, we consider a series of robustness checks and placebo exercises. First, due to the scarcity of alerts being issued on days with low temperatures, we restricted our main sample to region-days with a WBGT above 28.0$°C$. We test the robustness of our results in Appendix Table A7 by estimating our main specification across different sample cutoffs in WBGT. Across all sample considerations, the estimated coefficients remain consistent and statistically significant at the 1%-level. Furthermore, results remain unchanged if we do not restrict the sample to region-days with heat stroke counts below the 99th percentile (see Appendix Table A8).

As an alternative econometric robustness check, particularly given a relatively low number of region clusters ($n = 44$), we conduct a randomization inference test (Bertrand et al., 2004). To do so, we randomly reassign each region's treatment pattern (heat alerts) across regions, and estimate our main specification. We repeat this procedure 1,000 times, then plot the resulting placebo distribution of treatment effects alongside the actual treatment effect estimated in column (6) from Table 2. The results from this exercise are displayed in Appendix Figure A8, and once again we see that the primary treatment effect is highly statistically significant.

Next, we consider whether the source of the alert (previous-day 5pm forecast vs. same-day 5am forecast) had any differential bearing on heat strokes. In Appendix Table A9, we estimate our main specification

[43] For a detailed description of the data, estimation strategy, validation exercise, and results by age group, see the online appendix.

separately for alerts issued on the same day (5am) and the previous day (5pm). Note that an alert from 5pm the previous day was never "retracted" if the 5am same day WBGT forecast was below $33.0°C$. Therefore, the model for the same-day alert corresponds to our main model with an alert at any time. Meanwhile, only 4.6% of all region-days had an alert issued at 5am even though there was no alert from 5pm the previous day. In columns (1) to (4), we find that there is virtually no difference if we code the alert as stemming from 5am or 5pm. If we allow for interactions of both alerts, we find that around two thirds of the effect is driven by the same-day alert and only one third stems from the previous-day's alert that is confirmed on the same day.

Furthermore, in Appendix Table A10, we estimate our main specification for each year separately to test whether our results are driven by a specific year. Although the effect is strongest for 2020, it is very similar for the years 2021 to 2022, suggesting similar responses to heat alerts over time since the full roll out of the policy, and therefore similar effectiveness of the policy.

Because the introduction of the heat warning system coincided with the COVID-19 pandemic, it is possible that some of our results are due to policies affecting people's daily lives that coincided with heat alerts (e.g. movement restrictions). As a measure to prevent the spread of the virus, Japan's Prime Minister Shinzo Abe declared a state of emergency for Tokyo and six other regions on April 7, 2020. Although this measure was regional and temporary, it was extended to the entire country on April 16 and lasted until May 25. In the following two years, the Japanese government proclaimed four more states of emergency for various regions.[44] Although this measure did not restrict people's movement directly, people were asked to stay at home, and businesses as well as public facilities such as schools reduced operating hours or shut down completely. To test whether those measures coincided with heat alerts and thus drive our results, in Appendix Table A11, we estimate model (3) and include a dummy of whether a region-day was in a state of emergency. The results show that this policy did neither affect heat stroke nor the effect of a heat alert on heat stroke.[45]

In essence, our main specification corresponds to a two-way fixed effects regression in which units switch on treatment at different points in time for exactly one period. A recent literature has shown that in such staggered Difference-in-Differences (DiD) settings standard two-way fixed effects regressions are problematic in presence of treatment effect heterogeneity over time (Borusyak and Jaravel, 2018; De Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021). To address the issues arising from the negative weighting problem, in Appendix Table A12, we estimate our main specification using the alternative esti-

[44] In the online appendix, we present the timeline of these policies across prefectures.

[45] In the online appendix, we also show that the alerts did not have any effect on the number of Covid infections and that the effect of an alert on heat stroke is not driven by the number of Covid infections.

mator proposed by De Chaisemartin and d'Haultfoeuille (2020), which is robust to heterogeneous treatment effects. Coefficients are remarkably similar to our main results (Table 2, columns (1) to (4)) and highly statistically significant, leaving us confident that treatment effect heterogeneity is not a concern in our setting.

Finally, in Appendix Figure A9, we allow the effect to vary with subsamples of different WBGT cutoffs and for subsamples centered around specific WBGT values, respectively. Neither of these figures exhibit any heterogeneity. The results are rather stable around 0.15 to 0.18 and highly statistically significant. In contrast to the descriptive evidence from Figure 4, this suggests that the effect of a heat alert is the same for different temperatures. In particular, there is no differential effect for region-days with actual temperatures below and above the threshold of 33$^{\circ}C$.

7 Conclusion and Discussion

This study investigates the impacts of heat alerts on health and behavioral outcomes. The context of our study is Japan, which in 2020, introduced a comprehensive heat warning system to raise awareness of heat-related illnesses and promote heat stroke prevention measures. Our identification strategy utilizes variation across region-days in whether an alert was issued while flexibly controlling for same-day and previous-day wet bulb globe temperature (WBGT). We document substantial variation, with no manipulation, in whether an alert was issued for a region-day, controlling for observed WBGT - the lack of perfect collinearity comes from unobserved "forecasting errors," where alerts are determined solely by (inaccurate) forecasts.

Our findings suggest an increase in hospitalizations for heat stroke in response to a heat alert being issued. The effect exists across the spectrum of severity, including severe cases which require multiple days of hospitalization, and is most prevalent for adults and elderly. The effects are large, precisely estimated, and robust to a vast array of model considerations. Unsurprisingly, we further document that same-day and previous-day WGBT adversely impact heat stroke hospitalizations as well.

To unpack the potential mechanisms driving these results, we hypothesize that heat alerts can have various impacts on people's behavior. Primarily, we may expect some type of avoidance behavior, as one of the primary intentions of the program was to encourage people to engage in behaviors that mitigate the risk of heat - such a possible channel cannot solely explain our results however, given we observe a net increase in heat stroke hospitalizations. Thus, we consider an alternate set of potential mechanisms to explain the increase in heat stroke. First, alerts may raise one's awareness of the potential presence and risk of heat stroke (possibly without changing the actual incidence of heat stroke), leading to increased reporting. Second, we hypothesize that heat strokes may conflate with other sudden illnesses, such as cardiovascular disease and stroke, and when an alert is issued, there's effectively a "substitution" in how diseases are recorded, away from other causes and into heat stroke. Finally, an alert may be associated with other changing factors which then induce people to engage in behaviors that place them at a greater risk of heat stroke (e.g. leaving their home). In line with these potentially "adverse" responses, a previous literature from Japan has documented energy price sensitivity leading people to engage in some (unobserved) behaviors that increased overall mortality (Neidell et al., 2021; He and Tanaka, 2023).

To identify these potential behavioral responses, we utilize additional data from the population of ambulance records, Google Trends, and Google Community Mobility Reports. In total, we do not find any clear evidence for avoidance behavior except for an increase in the search term "air conditioner," which points toward people trying to combat high heat. Instead, we find that searches for various outdoor activities increased, and people were *more* likely to leave their home residence on heat alert days in favor of visiting both indoor and outdoor locations (i.e. parks). This evidence may still be consistent with avoidance behavior, particularly if individuals believe they are better protected from high heat away from home. This evidence may also partially reflect people leaving their residence to visit the hospital. We note, however, that the overall increase in leaving home residences (of over 11%) goes in direct contrast with the country's messaging encouraging people to stay at home. We still estimate an effect on overall ambulance transports for sudden and general illnesses, suggesting the increases in heat stroke are not driven by a substitution away from other diagnoses. We also estimate a significant increase in severe cases of heat stroke, which suggests the effects cannot be solely attributed to increased reporting of what were otherwise unidentified cases of heat stroke (in the absence of an alert). Finally, several explorations into mortality fail to uncover any evidence that the heat alerts reduced mortality.

Identifying the net welfare impacts of the heat alert system is challenging. For instance, suppose our effects were solely driven by increased reporting (with no effect on the actual incidence of heat stroke): on one hand, people would now receive proper medical treatment to prevent potential long-term negative impacts from heat stroke; on the other hand, due to hospital capacity constraints (particularly during a pandemic), medical resources devoted to a heat stroke could potentially have been "better" served to some other medical condition. Additionally, it is possible that the behaviors we observe (leaving home residence for both indoor and outdoor locations) constitute avoidance behavior, and that "actual" heat stroke incidence is decreasing, while observed heat stroke hospitalization is increasing solely due to the increased awareness effect. In this case, one could see welfare benefits from the warning system.

Still, significant concern remains on whether the behaviors we observe truly constitute heat-mitigating, welfare-improving responses from Japanese households. First, people leaving their households directly contrasts the government recommendation to stay at home. Second, one would expect that any gains from avoidance behavior would manifest in several potential ways, which we fail to see in the data, including intertemporal effects (we see no effect of heat alerts on future heat stroke) and overall mortality. Finally, as documented in Neidell et al. (2021) and He and Tanaka (2023), Japanese households were sufficiently energy-price sensitive to curb their behaviors in a matter that led to increased mortality. In our context, it may be that the heat alerts also trigger a desire to save energy among Japanese households.[46] Though merely speculation, we posit that our observed behavioral effects are driven by both an "avoidance" effect (i.e. people are better protected from high heat outside their homes) and a desire to save energy on (what they perceive to be) high heat days.

Ultimately, more research is needed to understand how people react to alerts, whether such responses are welfare-improving, and how to better improve alert systems. Concerns over "adverse" responses to alert systems are not unique to this context as well. Perhaps most infamously, in the summer of 2023, Maui failed to activate their audible warning system during wildfires that killed over 100 people, with county officials citing a fear that people would respond inappropriately to the alert by moving toward the (inland) fires.[47] In the UK, an alert system carried through an audible 10-second mobile phone alert has critics concerned that it puts people's safety at risk, including by potentially distracting drivers.[48] A separate literature has investigated the effects of warning labels regarding the credibility of specific news articles (i.e. "fake news"); in one prominent study, Pennycook et al. (2020) find that such labels improved the perceived credibility of other "false" articles that failed to receive a label. Though these contexts differ in many ways, they highlight how behavioral responses may conflict with the intended effects of an alert to generate adverse results.

[46] So far as we are aware, data do not exist on daily-level energy consumption and saving behavior. Still, energy costs are at an all time high in Japan, and overall, local governments have been encouraging households to conserve energy. Source: The Japan Times, https://www.japantimes.co.jp/news/2022/09/15/business/economy-business/japan-energy-request-winter/, retrieved September 13, 2023.

[47] Source: The New York Times, https://www.nytimes.com/2023/09/03/us/maui-wildfires-emergency-alerts.html, retrieved September 13, 2023.

[48] Source: Sky News, https://news.sky.com/story/emergency-alert-test-millions-of-phone-users-in-uk-receive-message-and-loud-alarm-12864193, retrieved September 13, 2023.

References

AKESAKA, M. AND H. SHIGEOKA (2023): "“Invisible killer”: Seasonal allergies and accidents," NBER Working Paper 31593.

BALLESTER, J., M. QUIJAL-ZAMORANO, R. F. MÉNDEZ TURRUBIATES, F. PEGENAUTE, F. R. HERRMANN, J. M. ROBINE, X. BASAGAÑA, C. TONNE, J. M. ANTÓ, AND H. ACHEBAK (2023): "Heat-related mortality in Europe during the summer of 2022," *Nature Medicine*, 1–10.

BARRECA, A., K. CLAY, O. DESCHENES, M. GREENSTONE, AND J. S. SHAPIRO (2016): "Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century," *Journal of Political Economy*, 124, 105–159.

BENMARHNIA, T., L. SCHWARZ, A. NORI-SARMA, AND M. L. BELL (2019): "Quantifying the impact of changing the threshold of New York City heat emergency plan in reducing heat-related illnesses," *Environmental Research Letters*, 14, 114006.

BERTRAND, M., E. DUFLO, AND S. MULLAINATHAN (2004): "How much should we trust differences-in-differences estimates?" *Quarterly Journal of Economics*, 119, 249–275.

BORUSYAK, K. AND X. JARAVEL (2018): *Revisiting event study designs*, SSRN Scholarly Paper ID 2826228, Social Science Research Network, Rochester, NY.

CATTANEO, M. D., M. JANSSON, AND X. MA (2018): "Manipulation testing based on density discontinuity," *The Stata Journal*, 18, 234–261.

——— (2020): "Simple local polynomial density estimators," *Journal of the American Statistical Association*, 115, 1449–1455.

CHETTY, R., J. N. FRIEDMAN, N. HENDREN, M. STEPNER, ET AL. (2020): "Real-time economics: A new platform to track the impacts of COVID-19 on people, businesses, and communities using private sector data," NBER Working Paper 27431.

DE CHAISEMARTIN, C., X. D'HAULTFOEUILLE, AND Y. GUYONVARCH (2019): "DID_MULTIPLEGT: Stata module to estimate sharp Difference-in-Difference designs with multiple groups and periods," Tech. rep., Boston College Department of Economics.

DE CHAISEMARTIN, C. AND X. D'HAULTFOEUILLE (2020): "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 110, 2964–2996.

DESCHÊNES, O. AND M. GREENSTONE (2011): "Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US," *American Economic Journal: Applied Economics*, 3, 152–185.

DESCHENES, O. AND E. MORETTI (2009): "Extreme weather events, mortality, and migration," *Review of Economics and Statistics*, 91, 659–681.

D'IPPOLITI, D., P. MICHELOZZI, C. MARINO, F. DE'DONATO, B. MENNE, K. KATSOUYANNI, U. KIRCHMAYER, A. ANALITIS, M. MEDINA-RAMÓN, A. PALDY, ET AL. (2010): "The impact of heat waves on mortality in 9 European cities: Results from the EuroHEAT project," *Environmental Health*, 9, 1–9.

Dodla, V. B., G. C. Satyanarayana, and S. Desamsetti (2017): "Analysis and prediction of a catastrophic Indian coastal heat wave of 2015," *Natural Hazards*, 87, 395–414.

Ebi, K. L., A. Capon, P. Berry, C. Broderick, R. de Dear, G. Havenith, Y. Honda, R. S. Kovats, W. Ma, A. Malik, et al. (2021): "Hot weather and heat extremes: Health risks," *The Lancet*, 398, 698–708.

Fernández-Villaverde, J. and C. I. Jones (2020): "Macroeconomic outcomes and COVID-19: A progress report," Tech. rep., National Bureau of Economic Research.

Ferris, J. S. and D. A. Newburn (2017): "Wireless alerts for extreme weather and the impact on hazard mitigating behavior," *Journal of Environmental Economics and Management*, 82, 239–255.

Geruso, M. and D. Spears (2018): "Heat, humidity, and infant mortality in the developing world," NBER Working Paper 24870.

Goodman-Bacon, A. (2021): "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, 225, 254–277.

Gutteling, J. M., T. Terpstra, and J. H. Kerstholt (2018): "Citizens' adaptive or avoiding behavioral response to an emergency message on their mobile phone," *Journal of Risk Research*, 21, 1579–1591.

He, G. and T. Tanaka (2023): "Energy saving may kill: Evidence from the Fukushima nuclear accident," *American Economic Journal: Applied Economics*, 15, 377–414.

Heutel, G., N. H. Miller, and D. Molitor (2021): "Adaptation and the mortality effects of temperature across US climate regions," *Review of Economics and Statistics*, 103, 740–753.

Iizuka, T., K. Nishiyama, B. Chen, and K. Eggleston (2021): "False alarm? Estimating the marginal value of health signals," *Journal of Public Economics*, 195, 104368.

Kadooka, Y., A. Asai, A. Enzo, and T. Okita (2017): "Misuse of emergent healthcare in contemporary Japan," *BMC Emergency Medicine*, 17, 1–6.

Karaivanov, A., S. E. Lu, H. Shigeoka, C. Chen, and S. Pamplona (2021): "Face masks, public policies and slowing the spread of COVID-19: Evidence from Canada," *Journal of Health Economics*, 78, 102475.

Kim, H. B., S. A. Lee, and W. Lim (2019): "Knowing is not half the battle: Impacts of information from the National Health Screening Program in Korea," *Journal of Health Economics*, 65, 1–14.

Kotharkar, R. and A. Ghosh (2022): "Progress in extreme heat management and warning systems: A systematic review of heat-health action plans (1995–2020)," *Sustainable Cities and Society*, 76, 103487.

Li, T., C. Chen, and W. Cai (2022): "The global need for smart heat-health warning systems," *The Lancet*, 400, 1511–1512.

Martiello, M. A. and M. V. Giacchi (2010): "High temperatures and health outcomes: A review of the literature," *Scandinavian Journal of Public Health*, 38, 826–837.

Mendolia, S., O. Stavrunova, and O. Yerokhin (2021): "Determinants of the community mobility during the COVID-19 epidemic: The role of government regulations and information," *Journal of Economic Behavior & Organization*, 184, 199–231.

MICHELOZZI, P., G. ACCETTA, M. DE SARIO, D. D'IPPOLITI, C. MARINO, M. BACCINI, A. BIGGERI, H. R. ANDERSON, K. KATSOUYANNI, F. BALLESTER, ET AL. (2009): "High temperature and hospitalizations for cardiovascular and respiratory causes in 12 European cities," *American Journal of Respiratory and Critical Care Medicine*, 179, 383–389.

MULLINS, J. AND P. BHARADWAJ (2015): "Effects of short-term measures to Curb Air Pollution: Evidence from Santiago, Chile," *American Journal of Agricultural Economics*, 97, 1107–1134.

MULLINS, J. T. AND C. WHITE (2019): "Temperature and mental health: Evidence from the spectrum of mental health outcomes," *Journal of Health Economics*, 68, 102240.

——— (2020): "Can access to health care mitigate the effects of temperature on mortality?" *Journal of Public Economics*, 191, 104259.

NEIDELL, M. (2009): "Information, avoidance behavior, and health: The effect of ozone on asthma hospitalizations," *Journal of Human Resources*, 44, 450–478.

NEIDELL, M., S. UCHIDA, AND M. VERONESI (2021): "The unintended effects from halting nuclear power production: Evidence from Fukushima Daiichi accident," *Journal of Health Economics*, 79, 102507.

NOONAN, D. S. (2014): "Smoggy with a chance of altruism: The effects of ozone alerts on outdoor recreation and driving in Atlanta," *Policy Studies Journal*, 42, 122–145.

ONO, M. AND M. TONOUCHI (2014): "Estimation of wet-bulb globe temperature using generally measured meterological indices," *Japanese Journal of Biometeorology*, 50, 147–157.

PALECKI, M. A., S. A. CHANGNON, AND K. E. KUNKEL (2001): "The nature and impacts of the July 1999 heat wave in the Midwestern United States: Learning from the lessons of 1995," *Bulletin of the American Meteorological Society*, 82, 1353–1368.

PENNYCOOK, G., A. BEAR, E. T. COLLINS, AND D. G. RAND (2020): "The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings," *Management science*, 66, 4944–4957.

ROBINSON, A., J. LEHMANN, D. BARRIOPEDRO, S. RAHMSTORF, AND D. COUMOU (2021): "Increasing heat and rainfall extremes now far outside the historical climate," *npj Climate and Atmospheric Science*, 4, 45.

SABERIAN, S., A. HEYES, AND N. RIVERS (2017): "Alerts work! Air quality warnings and cycling," *Resource and Energy Economics*, 49, 165–185.

SHRADER, J. G., L. BAKKENSEN, AND D. LEMOINE (2023): "Fatal errors: The mortality value of accurate weather forecasts," NBER Working Paper 31361.

SULYOK, M. AND M. WALKER (2020): "Community movement and COVID-19: A global study using Google's Community Mobility Reports," *Epidemiology & Infection*, 148, e284.

VAIDYANATHAN, A., S. SAHA, A. M. VICEDO-CABRERA, A. GASPARRINI, N. ABDUREHMAN, R. JORDAN, M. HAWKINS, J. HESS, AND A. ELIXHAUSER (2019): "Assessment of extreme heat and hospitalizations to inform early warning systems," *Proceedings of the National Academy of Sciences*, 116, 5420–5427.

WARD, A. L. S. AND T. K. BEATTY (2016): "Who responds to air quality alerts?" *Environmental and Resource Economics*, 65, 487–511.

WEINBERGER, K. R., X. WU, S. SUN, K. R. SPANGLER, A. NORI-SARMA, J. SCHWARTZ, W. REQUIA, B. M. SABATH, D. BRAUN, A. ZANOBETTI, ET AL. (2021): "Heat warnings, mortality, and hospital admissions among older adults in the United States," *Environment International*, 157, 106834.

WEINBERGER, K. R., A. ZANOBETTI, J. SCHWARTZ, AND G. A. WELLENIUS (2018): "Effectiveness of National Weather Service heat alerts in preventing mortality in 20 US cities," *Environment International*, 116, 30–38.

YAGLOU, C., D. MINAED, ET AL. (1957): "Control of heat casualties at military training centers." *Arch. Indust. Health*, 16, 302–16.

ZIVIN, J. G. AND M. NEIDELL (2009): "Days of haze: Environmental information disclosure and intertemporal avoidance behavior," *Journal of Environmental Economics and Management*, 58, 119–128.

——— (2014): "Temperature and the allocation of time: Implications for climate change," *Journal of Labor Economics*, 32, 1–26.

A Appendix Tables and Figures

Table A1: FDMA Data – Summary Statistics

	(1) Obs.	(2) Mean	(3) SD	(4) Min	(5) Max
Panel A: Overall					
# of transports	16879	370.778	435.00	45.00	2985.00
# of mild cases	16879	340.873	411.06	36.00	2779.00
# of severe cases	16879	26.296	24.36	0.00	193.00
# of deaths	16879	3.610	3.80	0.00	30.00
Panel B: Cause for ambulance transport					
# of transports (Fire & Water & Natural disaster)	16879	0.526	3.29	0.00	362.00
# of mild cases (Fire & Water & Natural disaster)	16879	0.384	3.11	0.00	349.00
# of severe cases (Fire & Water & Natural disaster)	16879	0.087	0.37	0.00	20.00
# of deaths (Fire & Water & Natural disaster)	16879	0.055	0.25	0.00	4.00
# of transports (Traffic accident)	16879	25.461	28.38	0.00	185.00
# of mild cases (Traffic accident)	16879	24.410	27.70	0.00	182.00
# of severe cases (Traffic accident)	16879	0.960	1.24	0.00	11.00
# of deaths (Traffic accident)	16879	0.091	0.31	0.00	4.00
# of transports (Work & Sports)	16879	7.301	8.41	0.00	83.00
# of mild cases (Work & Sports)	16879	6.931	8.20	0.00	80.00
# of severe cases (Work & Sports)	16879	0.349	0.64	0.00	7.00
# of deaths (Work & Sports)	16879	0.022	0.15	0.00	3.00
# of transports (Sudden & General illness)	16879	302.861	370.20	30.00	2644.00
# of mild cases (Sudden & General illness)	16879	280.895	351.71	27.00	2474.00
# of severe cases (Sudden & General illness)	16879	18.848	18.27	0.00	157.00
# of deaths (Sudden & General illness)	16879	3.118	3.42	0.00	28.00
# of transports (Violence & Self-harm)	16879	3.874	5.72	0.00	45.00
# of mild cases (Violence & Self-harm)	16879	3.221	4.97	0.00	40.00
# of severe cases (Violence & Self-harm)	16879	0.416	0.88	0.00	9.00
# of deaths (Violence & Self-harm)	16879	0.237	0.53	0.00	5.00
# of transports (Hospital transfer)	16879	30.507	28.67	0.00	195.00
# of mild cases (Hospital transfer)	16879	24.858	24.59	0.00	154.00
# of severe cases (Hospital transfer)	16879	5.593	5.50	0.00	46.00
# of deaths (Hospital transfer)	16879	0.057	0.24	0.00	3.00

Source: Authors' calculations using data from Fire and Disaster Management Agency (FDMA). *Note*: This table presents descriptive statistics of all key variables used in the analysis. WBGT $\geq$ 28. Data for 2017-–2021.

Table A2: Google Trends – Summary Statistics

	(1) Obs.	(2) Mean	(3) SD	(4) Min	(5) Max
Heat strokes	10647	20.670	23.04	0.00	100.00
Heat stroke alert	10647	8.813	19.35	0.00	100.00
Weather	10647	51.952	19.25	11.00	100.00
Temperature	10647	22.998	21.20	0.00	100.00
Air conditioner	10647	35.990	22.02	0.00	100.00
Outdoor	10647	22.020	25.03	0.00	100.00
Sea bathing	10647	12.092	21.35	0.00	100.00
Park	10647	30.643	15.39	0.00	100.00
Indoor	10647	14.085	21.47	0.00	100.00
Cinema	10647	29.417	25.29	0.00	100.00
Karaoke	10647	39.863	22.14	0.00	100.00

Source: Authors' calculations using data from Google Trends. *Note*: This table presents descriptive statistics of all key variables used in the analysis. Google Trends Searches for 熱中症 (heat stroke), 熱中症警戒アラート (heat stroke alert), 天気 (weather), 気温 (atmospheric temperature), エアコン (air conditioner), 屋外 (outdoor), 海水浴 (sea bathing), 公園 (park), 屋内 (indoor), 映画館 (cinema), and カラオケ (karaoke). WBGT $\geq$ 28. Data for 2020-–2022.

Table A3: Google Community Mobility Reports – Summary Statistics

	(1) Obs.	(2) Mean	(3) SD	(4) Min	(5) Max
Retail & recreation	10647	−6.569	9.59	−89.00	56.00
Grocery & pharmacy	10647	4.847	7.06	−82.00	45.00
Parks	10293	3.067	27.74	−80.00	280.00
Transit stations	10636	−15.544	15.29	−85.00	105.00
Workplaces	10647	−11.612	12.97	−78.00	8.00
Residential	10647	5.876	3.80	−2.00	34.00

Source: Authors' calculations using data from Google COVID-19 Community Mobility Reports. *Note*: This table presents descriptive statistics of all key variables used in the analysis. WBGT $\geq$ 28. Data for 2020-–2022.

Table A4: Effect of Heat Alerts on Heat Strokes by Place of Incidence

	# of heat strokes							
	(1) Home	(2) Workpl. (constr.)	(3) Workpl. (field)	(4) Educ.	(5) Public (ins.)	(6) Public (outs.)	(7) Street	(8) Other
$=1$ if heat alert	0.154***	0.204***	0.145**	0.208***	0.127***	0.195***	0.164***	0.167***
	(0.020)	(0.034)	(0.066)	(0.051)	(0.046)	(0.033)	(0.030)	(0.045)
WBGT	✓	✓	✓	✓	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓	✓	✓
Region-spec. trend	✓	✓	✓	✓	✓	✓	✓	✓
Mean heat strokes	5.672	1.544	.347	.809	1.090	1.722	2.128	.787
Observations	20612	20612	20612	20612	20612	20612	20612	20612

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from negative binomial regressions of the number of heat strokes by place of incidence in a region on the presence of a heat alert and a set of control variables. The mean number of heat strokes is for regions without heat alerts (control group). WBGT ≥ 28. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

Table A5: Effect of Consecutive Heat Alerts on Heat Strokes

	# of heat strokes			
	(1)	(2)	(3)	(4)
Heat strokes: total				
$= 1$ if heat alert$_{t-2}$			-0.004	-0.004
			(0.030)	(0.030)
$= 1$ if heat alert$_{t-1}$	0.016	0.004	0.005	0.005
	(0.015)	(0.021)	(0.021)	(0.021)
$= 1$ if heat alert$_t$	0.166***	0.156***	0.158***	0.166***
	(0.015)	(0.018)	(0.019)	(0.021)
Alert$_t$ $\times$ alert$_{t-2}$			-0.006	-0.051
			(0.041)	(0.057)
Alert$_t$ $\times$ alert$_{t-1}$		0.027	0.030	0.009
		(0.033)	(0.033)	(0.037)
Alert$_t$ $\times$ alert$_{t-1}$ $\times$ alert$_{t-2}$				0.064
				(0.051)
WBGT	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓
Region FE	✓	✓	✓	✓
Day FE	✓	✓	✓	✓
Region-spec. trend	✓	✓	✓	✓
Mean heat strokes	14.109	14.109	14.109	14.109
Observations	20600	20600	20600	20600

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from negative binomial regressions of the number of heat strokes in a region on the presence of a heat alert and a set of control variables. The mean number of heat strokes is for regions without heat alerts (control group). WBGT ≥ 28. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

Table A6: Effect of Heat Alerts on Heat Strokes on Weekdays vs. Weekends

	# of heat strokes			
	Weekdays		Weekends	
	(1)	(2)	(3)	(4)
$= 1$ if heat alert	0.171***	0.156***	0.180***	0.212***
	(0.018)	(0.018)	(0.029)	(0.030)
Mean WBGT	✓	✓	✓	✓
Lagged Mean WBGT	✓	✓	✓	✓
Region FE	✓	✓	✓	✓
Day FE	✓	✓	✓	✓
Region-spec. trend		✓		✓
Mean heat strokes	13.805	13.805	14.845	14.845
Observations	14791	14791	5821	5821

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from negative binomial regressions of the number of heat strokes in a region on the presence of a heat alert and a set of control variables. The mean number of heat strokes is for regions without heat alerts (control group). WBGT ≥ 28. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

Table A7: Robustness Check – Different Cutoffs for WBGT

	# of heat strokes					
	(1)	(2)	(3)	(4)	(5)	(6)
	WBGT $=$ $[20, max]$	WBGT $=$ $[25, max]$	WBGT $=$ $[28, max]$	WBGT $=$ $[20, 35]$	WBGT $=$ $[25, 35]$	WBGT $=$ $[28, 35]$
$= 1$ if heat alert	0.164***	0.164***	0.172***	0.165***	0.165***	0.173***
	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
WBGT	✓	✓	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓
Region-spec. trend	✓	✓	✓	✓	✓	✓
Mean heat strokes	8.432	10.369	14.100	8.227	10.126	13.788
Observations	37212	29708	20612	36919	29415	20319

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from negative binomial regressions of the number of heat strokes in a region on the presence of a heat alert and a set of control variables. The mean number of heat strokes is for regions without heat alerts (control group). Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

Table A8: Robustness Check – Effect of Heat Alerts on Heat Strokes, Full Sample

	Log (# of heat strokes +1)		# of heat strokes			
	OLS				Neg. binomial regression	
	(1)	(2)	(3)	(4)	(5)	(6)
= 1 if heat alert	0.213***	0.213***	6.415**	6.702**	0.191***	0.190***
	(0.021)	(0.020)	(2.641)	(2.769)	(0.021)	(0.020)
WBGT	✓	✓	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓
Region-spec. trend		✓		✓		✓
Mean heat strokes	15.772	15.772	15.772	15.772	15.772	15.772
Observations	20816	20816	20816	20816	20816	20816

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from OLS and negative binomial regressions of the number of heat strokes in a region on the presence of a heat alert and a set of control variables. Data is *not* restricted to the 99^{th} percentile of the number of heat strokes. The mean number of heat strokes is for regions without heat alerts (control group). WBGT $\geq$ 28. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

Table A9: Robustness Check – Alert Due to Same-Day 5am vs. Previous-Day 5pm Forecasts

	# of heat strokes					
	Morning		Night		Morning & night	
	(1)	(2)	(3)	(4)	(5)	(6)
= 1 if heat alert at 5pm			0.171***	0.167***	0.061***	0.061***
			(0.017)	(0.015)	(0.020)	(0.018)
= 1 if heat alert at 5am	0.174***	0.172***			0.132***	0.130***
	(0.016)	(0.015)			(0.020)	(0.020)
WBGT	✓	✓	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓
Region-spec. trend		✓		✓		✓
Mean heat strokes	14.100	14.100	14.326	14.326	14.326	14.326
Observations	20612	20612	20612	20612	20612	20612

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from negative binomial regressions of the number of heat strokes in a region on the presence of a heat alert and a set of control variables. The mean number of heat strokes is for regions without heat alerts (control group). WBGT $\geq$ 28. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

Table A10: Robustness Check – Effects by Year

	# of heat strokes		
	2020	2021	2022
	(1)	(2)	(3)
$=1$ if heat alert	0.239***	0.162***	0.179***
	(0.085)	(0.032)	(0.021)
WBGT	✓	✓	✓
Lagged WBGT	✓	✓	✓
Region FE	✓	✓	✓
Day FE	✓	✓	✓
Mean heat strokes	13.487	9.782	11.949
Observations	3500	3241	3906

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from negative binomial regressions of the number of heat strokes in a region on the presence of a heat alert and a set of control variables. The mean number of heat strokes is for regions without heat alerts (control group). WBGT ≥ 28. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

Table A11: Robustness Check – State of Emergency (COVID-19)

	# of heat strokes			
	(1)	(2)	(3)	(4)
$=1$ if heat alert	0.175***	0.173***	0.177***	0.176***
	(0.016)	(0.015)	(0.017)	(0.016)
$=1$ if state of emerg.	0.018	0.041	0.023	0.046
	(0.049)	(0.056)	(0.050)	(0.056)
Alert $\times$ state of emerg.			-0.025	-0.027
			(0.043)	(0.044)
WBGT	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓
Region FE	✓	✓	✓	✓
Day FE	✓	✓	✓	✓
Region-spec. trend		✓		✓
Mean heat strokes	14.100	14.100	14.100	14.100
Observations	20612	20612	20612	20612

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from negative binomial regressions of the number of heat strokes in a region on the presence of a heat alert, presence of state of emergency, and a set of control variables. The mean number of heat strokes is for regions without heat alerts (control group). WBGT ≥ 28. Robust standard errors clustered at the regional level are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively.

Table A12: Robustness Check – Estimator of De Chaisemartin and d'Haultfoeuille (2020)

	Log(# of heat strokes +1)		# of heat strokes	
	(1)	(2)	(3)	(4)
$= 1$ if heat alert	0.160***	0.160***	3.799***	3.798***
	(0.028)	(0.028)	(0.935)	(0.935)
WBGT	✓	✓	✓	✓
Lagged WBGT	✓	✓	✓	✓
Region FE	✓	✓	✓	✓
Day FE	✓	✓	✓	✓
Region-spec. trend		✓		✓
Mean heat strokes	14.100	14.100	14.100	14.100
Observations	20612	20612	20612	20612

Source: Authors' calculations using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This table shows results from OLS regressions of the number of heat strokes in a region on the presence of a heat alert and a set of control variables, using the estimator of De Chaisemartin and d'Haultfoeuille (2020). WBGT and lagged WBGT is controlled for linearly. The mean number of heat strokes is for regions without heat alerts (control group). WBGT ≥ 28. Robust standard errors clustered at the regional level are obtained from 199 bootstrap replications and are given in parentheses. *, **, and *** denote significance at the 10%-, 5%-, and 1%-level, respectively. Results are obtained from the Stata command `did_multiplegt` of De Chaisemartin et al. (2019).

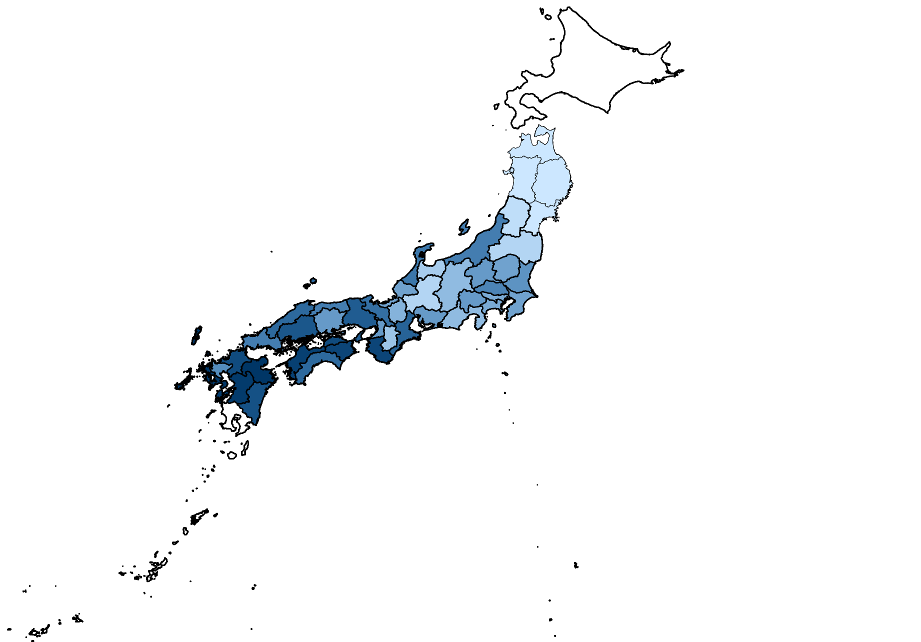


Figure A1: Share of Dates with Heat Alert in 2022

Source: Authors' presentation using data from Ministry of Environment. *Note*: This graph shows the share of dates with a heat alert between May 1 and September 30, 2022, across Japan. Regions in white are dropped due to imperfect mapping of regions to prefectures (described further in text). Regions in bright blue had zero alerts. Darker color indicates a higher share.

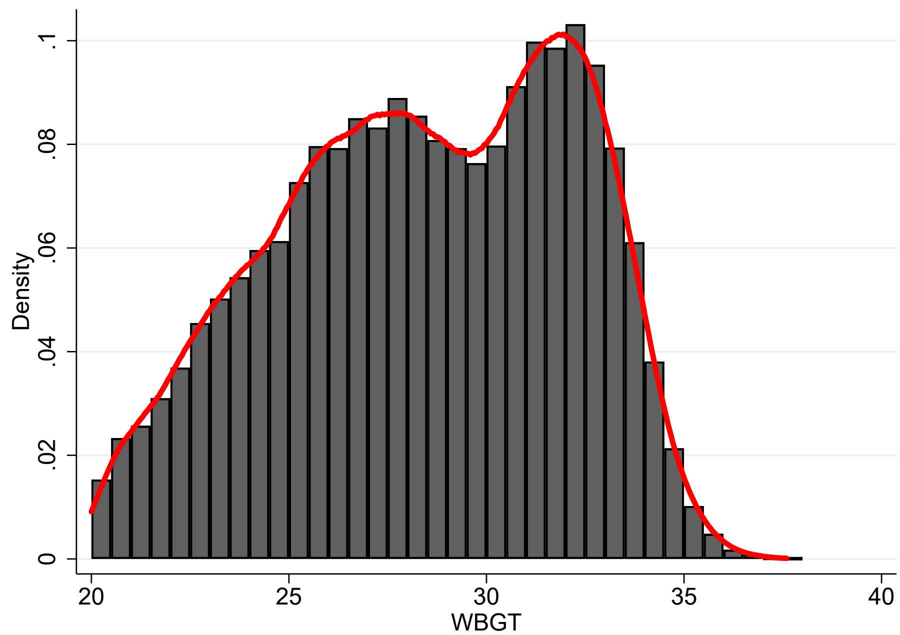
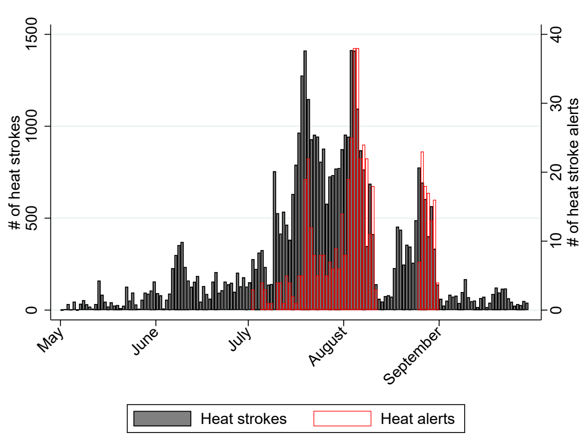
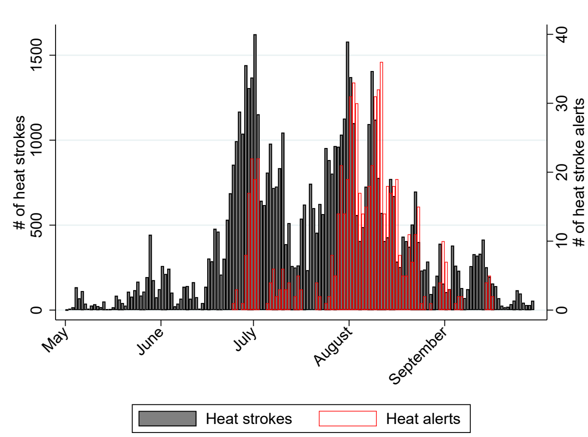


Figure A2: Distribution of Realized WBGT

Source: Authors' presentation using data from Ministry of Environment. *Note*: This graph shows the distribution of maximum realized WBGT in a region for the dates between May 1 and September 30 for the years 2017 to 2022. A kernel density estimate based on the epanechnikov kernel and the rule-of-thumb bandwidth is shown in red.



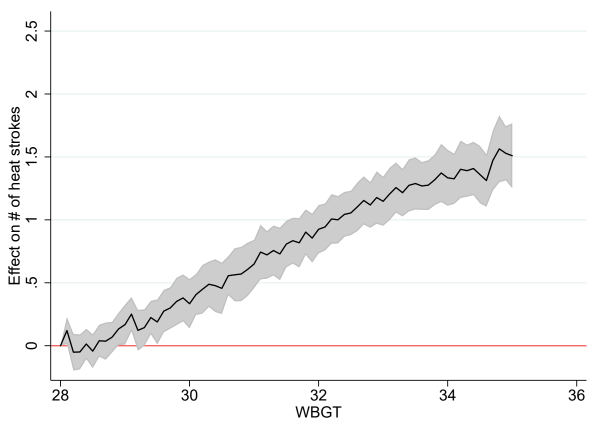
(a) 2021



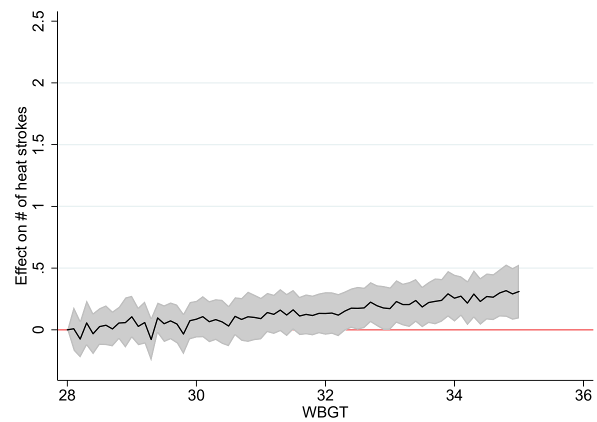
(b) 2022

Figure A3: Heat Strokes and Heat Alerts Over Time

Source: Authors' presentation using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: These graphs show the distribution of the number of heat strokes and heat alerts for 2021 (a) and 2022 (b).



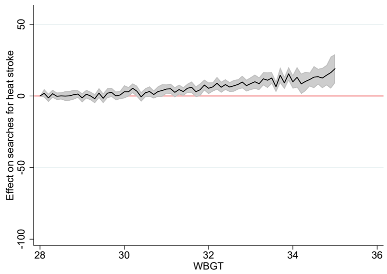
(a) Actual WBGT



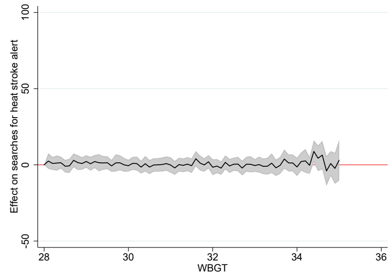
(b) Lagged WBGT

Figure A4: Relationship between WBGT and Heat Strokes

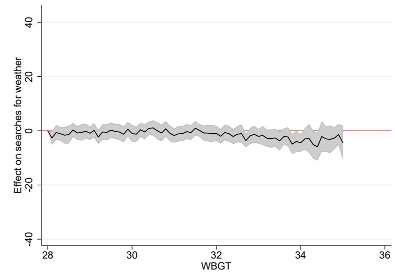
Source: Authors' presentation using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: These graphs show the relationship between the WBGT index (a) and the lagged WBGT index (b), respectively, and heat strokes according to our baseline specification presented in column (6) of Table 2. The estimated coefficients are presented for WBGT $= [28, 35]$, whereas WBGT $= 28.0$ serves as baseline. The grey area denotes 95%-confidence intervals based on robust standard errors clustered at the regional level.



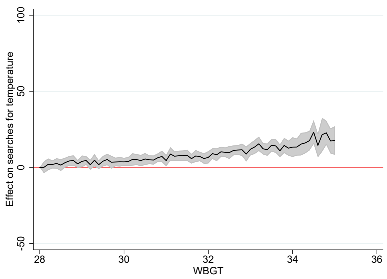
(a) “Heat stroke”



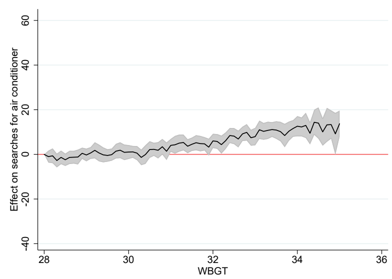
(b) “Heat stroke alert”



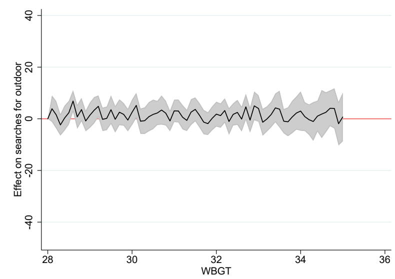
(c) “Weather”



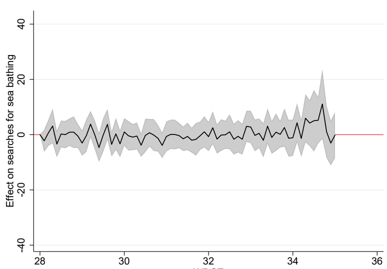
(d) “Temperature”



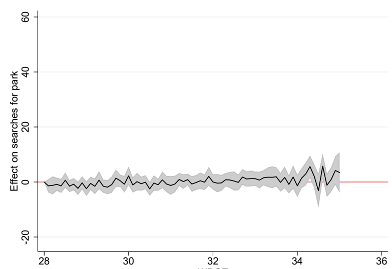
(e) “Air conditioner”



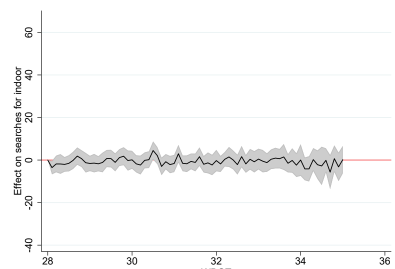
(f) “Outdoor”



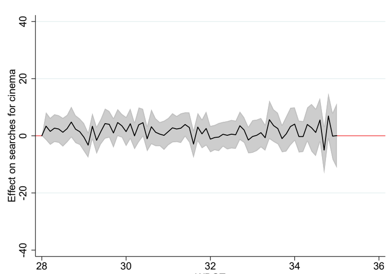
(g) “Sea bathing”



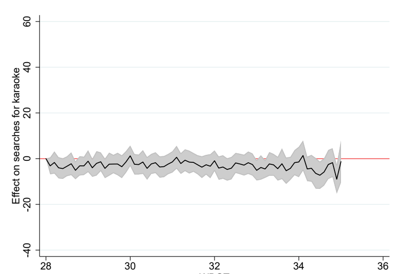
(h) “Park”



(i) “Indoor”



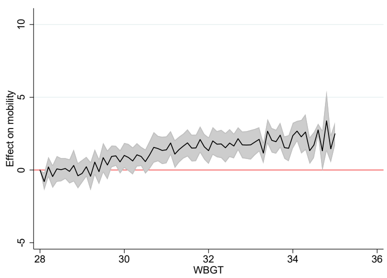
(j) “Cinema”



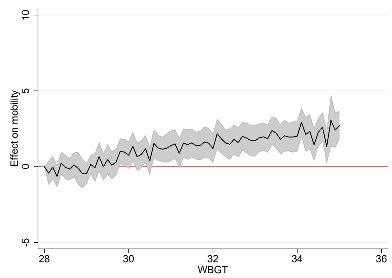
(k) “Karaoke”

Figure A5: Relationship between WBGT and Google Trend Searches

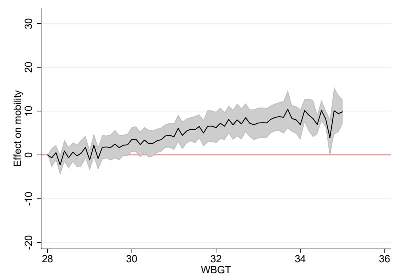
Source: Authors’ presentation using data from Google Trends and Ministry of Environment. *Note*: These graphs show the relationship between the WBGT index and Google Trend Searches for 熱中症 (heat stroke), 熱中症警戒アラート (heat stroke alert), 天気 (weather), 気温 (atmospheric temperature), エアコン (air conditioner), 屋外 (outdoor), 海水浴 (sea bathing), 公園 (park), 屋内 (indoor), 映画館 (cinema), and カラオケ (karaoke), respectively, according to our baseline specification presented in Table 7 and 8. The effects are presented for WBGT $= [28, 35]$, whereas WBGT $= 28.0$ serves as baseline. Data for 2020–2022. The grey area denotes 95%-confidence intervals based on robust standard errors clustered at the regional level.



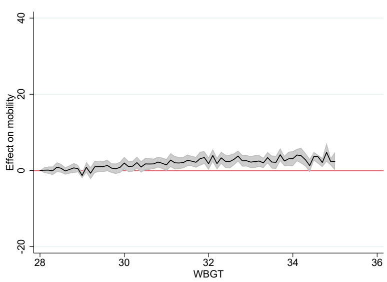
(a) Retail & recreation



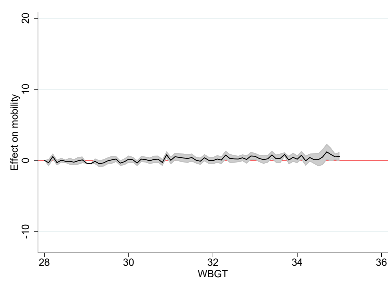
(b) Grocery & pharmacy



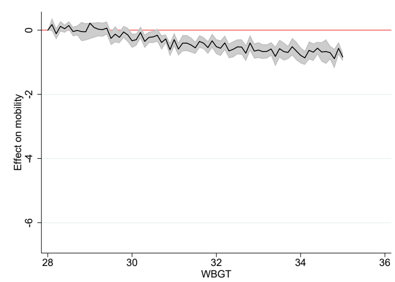
(c) Parks



(d) Transit stations



(e) Workplaces



(f) Residential

Figure A6: Relationship between WBGT and Mobility

Source: Authors' presentation using data from Google COVID-19 Community Mobility Reports and Ministry of Environment. *Note*: These graphs show the relationship between the WBGT index and the percentage changes in number of visits to specific places in a region, according to our baseline specification presented in Table 9. The effects are presented for WBGT $= [28, 35]$, whereas WBGT $= 28.0$ serves as baseline. Data for 2020–2022. The grey area denotes 95%-confidence intervals based on robust standard errors clustered at the regional level.

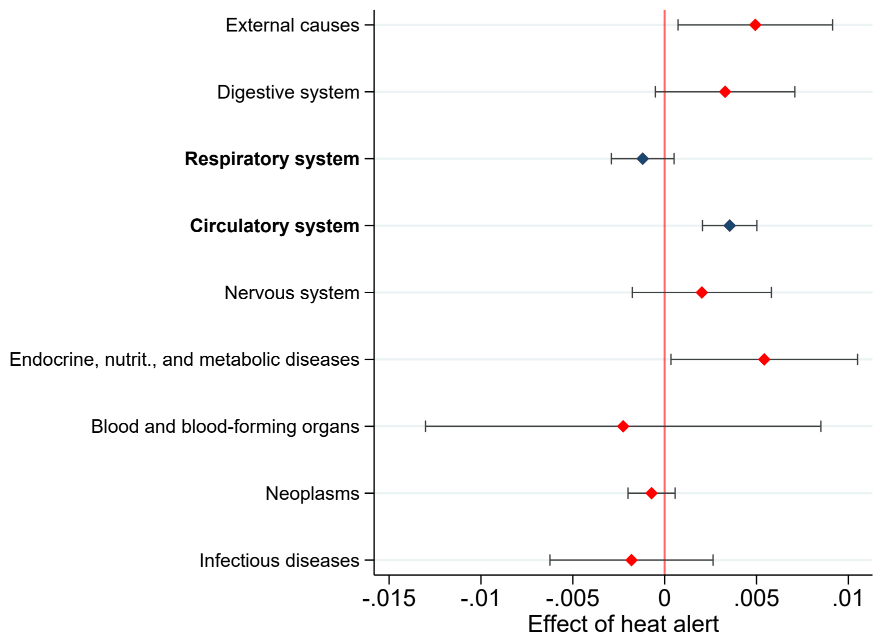


Figure A7: Effect of Heat Alerts on Cause-Specific Mortality

Source: Authors' presentation using data from Ministry of Environment, Ministry of Health, Labor, and Welfare, and Fire and Disaster Management Agency (FDMA). *Note*: This graph shows the effects of the number of deaths on the number of heat alerts in a region and a set of control variables using negative binomial regressions. Observations are at the region-month level and we estimate the following model:

$$Deaths_{crym} = \alpha + \beta Alert_{rym} + Share\_Temp_{rym}^{k} + \delta_{ry} + \lambda_{rm} + \theta_{ym} + \epsilon_{crym},$$

where $Deaths_{crym}$ is the number of deaths in category c in region r in year y in month m. $Alert_{rym}$ is the number of heat alerts, $Share\_Temp_{rym}^{k}$ is the share of days in a month with WBGT in the interval k (i.e. $[20; 21)$, ..., $[34; max]$), and δ_{ry}, λ_{rm}, and θ_{ym} are region-by-year fixed effects, region-by-month fixed effects, and year-by-month fixed effects, respectively. ϵ_{crym} is an error term. In the online appendix, we present results by age group. $N = 1276$. 95%-confidence intervals are based on robust standard errors clustered at the regional level.

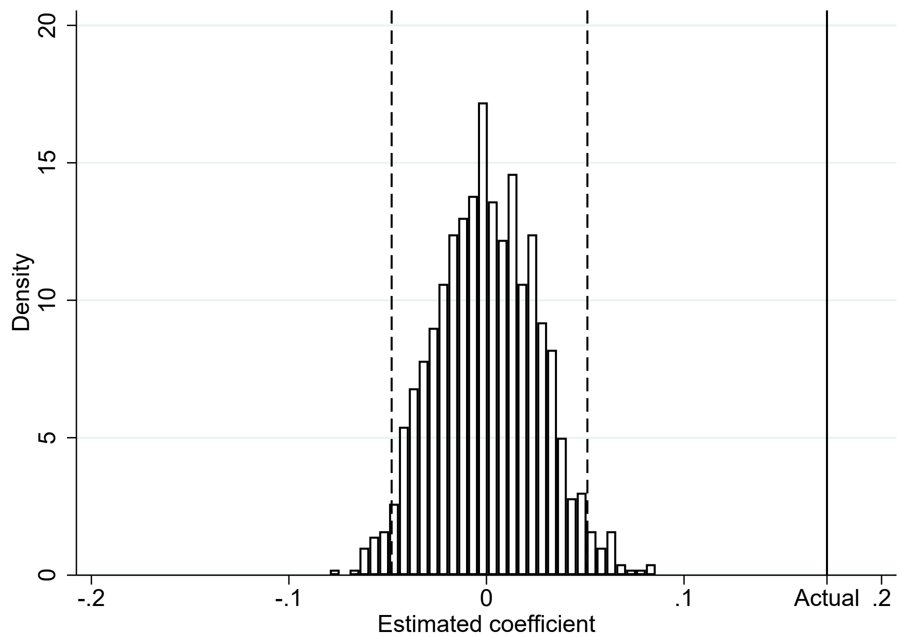
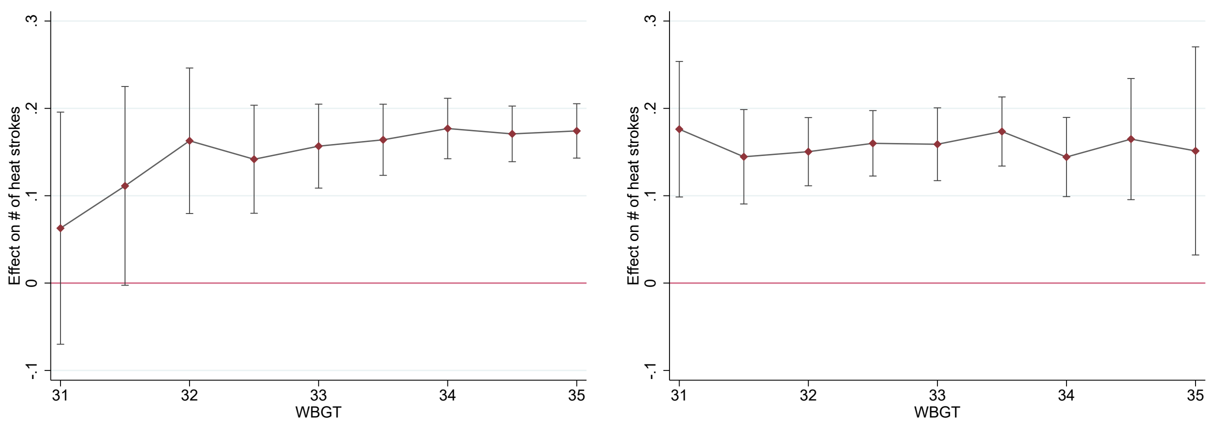


Figure A8: Placebo Distribution of Heat Alert Effects

Source: Authors' presentation using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: This graph shows the distribution of placebo negative binomial regressions of the number of heat strokes in a region on the presence of a heat alert and a set of control variables. The actual estimate based on column (6) of Table 3 is shown as solid line. Heat alert patterns over time are randomly assigned to regions (without replacement) and the negative binomial regression as described before is run. This procedure is repeated 1,000 times. This approach is based on Bertrand et al. (2004). Dashed lines indicate empirical 95%-confidence intervals.



(a) Increasing sample size

(b) Moving average

Figure A9: Effect of Heat Stroke Alerts on Heat Strokes by WBGT Intervals

Source: Authors' presentation using data from Ministry of Environment and Fire and Disaster Management Agency (FDMA). *Note*: These graphs show the effect of heat stroke alerts on heat strokes according to our baseline specification, while increasing the sample size in 0.5-intervals of the WBGT index (a) and using subsamples for specific WBGT intervals, respectively. In panel (a), presented estimates are the effect of an alert for the sample with temperatures just including the respective WBGT. In panel (b), presented estimates are the effect of an alert for a sample with temperatures centered around the respective WBGT $\pm$ 1. WBGT $\geq$ 28. 95%-confidence intervals are based on robust standard errors clustered at the regional level.